New Jersey Department of Community Affairs
Patricia Q. Sheehan, Commissioner

Division of State and Regional Planning
Richard A. Ginman, Director

Bureau of Regional Planning

A COASTAL AREA INFORMATION SYSTEM FOR THE NEW JERSEY DEPARTMENT OF ENVIRONMENTAL PROTECTION: A FEASIBILITY STUDY

by: Dennis K. Jones,
staff director
Robert Rockwell,
William K. Power, Jr.

August 1975
revised September 1975

Prepared for:
State of New Jersey
Department of Environmental Protection
Office of Environmental Analysis
Coastal Zone Management Program

Funded Under Federal Grant No. 04-4-158-50028
Office of Coastal Zone Management
National Oceanic and Atmospheric Administration
U. S. Department of Commerce

TABLE OF CONTENTS

TABLE OF CONTENTS--APPENDIXES

I. THE SYSTEM

The Department of Environmental Protection is charged by both State law[1] (hereafter called CAFRA) and Federal law[2] with overseeing economic development in the New Jersey coastal area. This regulation must take place at the level of the individual proposal for development, and the evaluation is to be in terms of local and regional criteria derived from comprehensive resource planning.[3] Planning policy must take into account both the expressed values of those who live in the coastal area as well as the reasonable expectations of those who would make their living there.[4]

Neither permit review, nor regional planning, nor public involvement are novel requirements in themselves. What is new, and what gives them power, is the mandate that these activities shall be coordinated expressions of a single policy.

Land use in the Coastal Area is the spatial "solution" to a complex series of private, public, economic, social, and environmental "equations". The objective is the realization of self-sustaining and long-lasting land use patterns in New Jersey's coastal area. The Division of State and Regional Planning, under the Department of Environmental Protection's direction, has looked at the overall mechanism for achieving the objectives of CAFRA and the Federal Coastal Zone Management Act. Although the Department of Environmental Protection does not have the power to dictate land use directly, it does have powers to plan, to regulate land use, and to interact with the public. The power of planning is in its ability to shape policy and provide analytic tools for the review process. The power of regulation is in its ability to exercise selectivity over proposed development. The power of public interaction is in promoting wider appreciation of environmental issues and helping developers design projects more sensitive to environmental concerns.

The greatest amount of work, thus far, has gone into permit review. This is because permit decisions can not wait and must be made. But it should be realized that permit review only relates to the fractional amount of the coastal area being considered for development and subject to Departmental review. The planning power is more important because it can directly affect policy and analysis. Public information, however, is potentially the Department's most effective power. In the long run, it can touch all citizens, influence their behaviour, provide better knowledge of all the coastal area, and give insight and guidance to land use decisions whether subject to review or not.

The Department's environmental management task cannot be taken too seriously. Between State legislation, federal legislation, and strong administrative and judicial rulings DEP has the mandate to "do something". Federal funding gives DEP the resources to do something significant. The CAFRA legislated timetable puts pressure on DEP to do something fast.

DSEP concentrated attention on the recurring needs for information, analysis, and documentation that characterize DEP planning, permit, and public information activities. The overlapping information needs of the public, the planning staff, and the permit review staff constitute a program for a "system". We are proposing that DEP adopt a comprehensive solution to this program. We are furthermore proposing major methodological elements of the system solution. The methodology is based on time-saving techniques for assembling information, doing analysis, evaluating permit applications, and documenting decisions. We believe that the threshold capability for any such DEP information system would be that it constitute a medium for promulgating comprehensive policy, that it be accessible to the public, that it increase the power and quality of analysis, that it improve permit application "through-put", and that it be more cost-effective than present methods.

The Department of Environmental Protection with the help of the Departments of Treasury and Transportation can specify such a system from resources at the Transportation Computer Center, equipment and technique extensions available from outside vendors, and staff already in the Department of Environmental Protection. The implementation of a system is feasible within the legislated timetable given normal budgetary and bureaucratic constraints and the anticipated level of federal funding.

---

[1]New Jersey's Coastal Area Facility Review Act of 1973.

[2]Coastal Zone Management Act of 1972.

[3]Coastal Area Review Board, Decision on Appeal of Lehigh Construction Company Condiminium, Toms River, January 1975.

[4]"Mount Laurel Decision" of the New Jersey Supreme Court, March 1975.

II. INFORMATION NEEDS

The Department of Environmental Protection has the power to shape the future of the coastal area -- not just by protecting the environment, but by establishing sound and self-sustaining patterns of land use. The skillful exercise of its authority and responsibility will make exacting demands on the Department for information, analysis and documentation.

Public Participation

The Department acts as the agency for the public's interest in the coastal area. This is a central fact that underlies both the mandate and the specific activities of the Department. The public has rights to information that the Department is obliged to recognize. One reasonable expectation is that the concerned private citizen should have ready access to facts in the public domain. Another reasonable expectation is that the planning process should be open to review and comment. There should also be a means to provide the potential developer with design assistance in advance of a formal permit request. Permit review policy should derive from planning models to which the public has made inputs. Permit decisions and opinions should be documented in a way that non-technical persons can easily comprehend.

Meaningful public participation depends wholly on understandable media of communication. Crossing the gap between the trained analyst and the concerned but non-specialist citizen is where many attempts at public involvement have foundered. Graphic display techniques, especially when computerized, combine informational fidelity with ease of comprehension. We think this is an essential public information requirement. Public access to information must be coupled with understanding.

Planning

Planning is the simultaneous understanding of many factors in one space. Planning is both process and product.

From the viewpoint of process, planning is analysis. It should be comprehensive, which is to say it should be a rigorous and dispassionate evaluation of a broad spectrum of data. It should have a capability to integrate current data. It must be able to cope with large volumes of data in a wide range of formats.

Planning analysis should also be able to utilize complex information when the interrelations are known. It should also have a means for equilibrating multiple factors, and for delineating regions or zones on the basis of specified factors.

From the viewpoint of product, planning is display. It is the representation of the spatial conclusions emerging from the planning analysis process. Quality of content is only one part of the planning product. The communicability is of equal importance. If planning is to have an input to permit review, it must tailor itself to decision-making at the level of the individual site. If planning is to lead policy, it must be versatile enough to illustrate impacts and ramifications. If planning is to benefit from public participation, a convenient and compatible medium of communication must be found. If planning is to be beneficial to public users, it must be understandable and straightforward.

Again, a versatile graphic display capability seems necessary. In this case, it is the product-producing extension of the analytic planning process.

Permit Review

Permit review operations are also a combination of process and product. The process part involves the certification of an application's completeness and an analysis of its strengths and weaknesses. The product part of permit review is the decision either approving or denying a permit. Present Department of Environmental Protection operations might be called long in process and short in product.

The information requirements of a permit system are many and complex. We estimate that an average permit application receives approximately seventeen separate functional reviews from ten reviewing agencies. Review personnel assemble data relevant to an application, analyze it according to various review criteria, and present conclusions. The review process, then, though at a smaller geographic scale than planning, is in a large way duplicative of planning information activities. From a system viewpoint, the planning and permit review processes could be built around a unified capability for the assembly, analysis, and display of geographic data.

The present permit process is characterized by a heavy workload and a slower than desired rate of output. When the State's economy inevitably turns upward, the rate of incoming applications will increase. This will strain the present process. Solution strategies could include hiring more staff, automating operations, eliminating redundancy, reducing delays, or forestalling incomplete applications.

The current premium attached to speedy review does not alter the importance of other objectives. The permit process must balance between prompt "throughput", high-quality analysis, fair treatment under equitable standards, and meaningful public participation.

The product of review is the permit decision which should be more than a summary approval or denial. The decision is the only expression of the review process visible to the public. As such it should manifest sound analysis, equitable criteria, fair and uniform treatment, public input, forthright deliberation, and timely and decisive action.

In the case of a permit denial, the developer has a reasonable right to learn how his project is deficient, the statutory right that DEP is exercising, the rights of appeal, citations of precedents for the action, and ideally a range of alternative locations where the projected land use would be more appropriate.

In the case of a conditional approval, the developer should be provided full details on the conditions attached to the permit, citations of precedent actions, possible design suggestions, and guidelines for compliance.

The information system requirements of permit review exceed those of planning. Permit review, because of its site-specificity, needs an order more detailed data. If the developer is to supply this level of data, the question of validation or authentication probably arises. There should also be an enforcement or land use monitoring procedure to insure that projects do not elude the permit process. There must be a straightforward means for relating site review back to comprehensive regional planning. All these information requirements can be met with a system. The Department of Environmental Protection can reasonably hope, not just to plan comprehensively, but also to be implementing comprehensive policy with each and every permit decision.

Project Design Assistance

The most prominent information need is that of the potential developer. To propose a land use change, he must start his information search from next to no data. DEP is in a position to do both the developer and itself a favor by extending the full resources of its planning and policy to interested parties. It may, on the one hand, forestall the weakest propositions. On the other hand, it stands to improve almost any reasonable proposal. The more the advance help that is extended to the developer before he applies to DEP for a permit, the better the chances of finding fully satisfactory solutions to problematic issues. Its a waste of good ideas to only air them during review.

In the true spirit of public information and participation, DEP could consider its role as being regulatory only where project design assistance and public information have failed.

The System: More Than the Sum of its Parts

How is it we can talk of a single program for a system that comprises such distinct activities as planning, permit review, public participation, and project design? The answer lies in the substantial duplication of information uses among the several activities. It is these common information processing requirements that consititute the program for the system. And it is the interrelation of these requirements that make it possible to design a central system capability as a unifying mechanism for four discrete parts of CAFRA.

The advantage of acknowledging the full range of information requirements at the outset is that the resulting system will do more than simply solve the specific needs of individual activities. It will also constitute a versatile resource available for other DEP uses outside of CAFRA. Since the system stands to do some things easier, some things faster, and other things better than present operations, the likelihood of other DEP uses and functions discovering the utility of the CAFRA information capability is great. A bonus reason for developing broad information capabilities in a CAFRA system is their use in DEP activities parallel to CAFRA functions but outside its specific mandate. Such capabilities as automatically mapping any geographic data and engaging in computerized graphic analysis of land-based issues would find many DEP users. (Since, for example, some of the techniques we propose would speed up CAFRA permit activities it might prove feasible to provide automated analysis assistance to other DEP permit activities.)

Capabilities developed and used on a limited basis in CAFRA can always be extended to a larger scope at a subsequent time. For example, once the system proves itself in coastal zone comprehensive planning, similar activity could be undertaken for other areas of the State. Once an accessible CAFRA public information and participation process is established, mechanisms might be forged for a statewide environmental protection public participation process. Once the techniques for helping developers propose and design better projects have been tested and demonstrated in the small geographic domain of the coastal area, voluntary assistance could be offered to the entire State.

Many such scenarios can be suggested. The common element is having a versatile central capability for geographic data processing, techniques for graphic display and analysis, and convenient means for the direct user interaction with the information and the operation of the system. Technique, once developed and implemented in a staff and machine system, is a great resource. Its application and coverage can be extended to additional territory or to additional operations at no additional cost. An operational coastal area information system (CAIS), once achieved, could be extended to include the entire State in its geographic base and other DEP activities in its applications.

III. METHODOLOGY

For the purposes of this report we have consistently distinguished between data and information. Data are uninterpreted facts; information is an aggregation of facts to which meaning is associated. A data processing system performs logical and quantitative operations with data. An information system, by our definition, assembles data for interpretation. Computers can process data, and people can interpret meaning. The degree to which these two processes can be made to work together is the measure of an information system. High information yield should be the goal of a CAFRA man-machine information system.

Conventional data processing systems have a low information yield. The reason for this is partly inadequate analytic technique, and partly the misguided nature of the data stock. This will be discussed thoroughly in subsequent parts of the report. The task of this section is to present expedients of data and of technique that, taken together, constitute a high-yield methodology for a geographic information system.

At the beginning of this study it was not apparent that the proposition of methodology would be so central to our work. It was believed, perhaps naively, that since the information requirements deducible from CAFRA only sketch the general structure of a system, the detailed program for a system would become evident in the daily activities of DEP planning and review agencies.

This Division conducted a series of technical interviews with the agencies that review CAFRA permit applications. We participated in these exchanges and tried to answer a long series of highly detailed information need and data use questions. We had hoped these technical interviews would provide some additional specificity for the general program. This did not work, but we did learn about DEP data on the one hand and about performance goals for data use on the other hand. The paramount question about conventional data processing and data base approaches is whether they have utility equivalent to the original and the recurring commitments of resources they cost.

It is typical in governmental planning and decision-making situations for actions and policy to be settled on a purely intuitive basis, and for data processing to engage in after-the-fact attempts to validate or justify hunches. This is backward. It is also backward to build a data base by first identifying data categories of apparent importance, and then collecting the highest possible grade of data in each category.

In both of the above typical situations, too little attention has gone to determining what information is actually wished from data, and too little planning has gone to how data can actually be used to produce information. It is now possible to simplify data needs, amplify the indwelling information content of data, and interactively explore intuitions and hunches in advance of decisions and policy. The connected series of techniques that we will propose constitute a methodology.

(1) Data Simplification

The first thrust of methodology should be to reduce data needs to convenient, reasonably manageable proportions. Principal strategies include a systematic philosophy of data "thrift", utilization of modelling techniques, the use of data surrogates, heavy reliance on self-updating forms of data, and a geographic data structure.

A "thrifty" philosophy of data means only storing and maintaining data that are necessary. This generally means avoiding the inclusion of data that "seem" important but are not known to be needed. The concept of a comprehensive inventory of data types only makes sense if there is a parallel comprehensive use of data.

The most prominent unnecessary class of data is the bulk of detailed data at the site-specific level that usually gluts most land-based information systems. In three of the four CAFRA information needs (planning, public participation, and project assistance), too much data can complicate matters and obfuscate real issues. Only in the permit review process is a great deal of detailed site-specific data essential. And even in this case, it is not necessary to store it for all time. The applicant developer can be expected to provide much if not all required data in the permit application. Given proper assistance, it is reasonable to provide all site-specific evaluative data as a single verified and validated file in machine-readable format. It need stay active in the system only long enough to enable processing. Thereafter, one need only retain those parts of the file that elucidate the regional file (in the case of a permit denial) or that have consequences on future decisions (in the case of a permit approval).

The second strategy for keeping data within manageable bounds is modelling. Techniques for using computers to model complex phenomena for the purposes of specific prediciitons are well-known, and need not be belabored here. Models are frequently used in conjunction with remote data, generally require large-capacity computer processing facilities, and are esteemed by virtue of their meaningful derivative information products rather than their complex use of raw input data.

A third strategy is to make creative use of data surrogates, or "indicators" as they are sometimes called. Using data surrogates means substituting available known data for related but unavailable data. Much depends on the extent of correspondence between the "indicator" and the phenomenon for which it is to substitute. Swamp maples may well indicate the presence of swamps (or swamps may be defined to mean the presence of swamp maples). Alluvial soils may be an indicator of flood-prone areas. Infant and neo-natal mortality may be an indicator of the general service level of a health care delivery system.

The intent is that the indicator data should substitute closely for a related type of data which is an important information dimension but either is difficult to obtain or is impossible to maintain. Initial studies have to ascertain correspondences. But once a correspondence has been validated, the surrogate data can be used. Thereafter, occasional sampling or spot-checking will be sufficient to guarantee prediction within statistically acceptable levels of confidence.

Remote sensing relies on the concept of data surrogates. Remote sensing from earth satellite has had extensive application in land use classification, and is also being used to monitor land use changes. In these situations, the satellite is not interpreting land use directly. Rather, it merely measures the spectral reflectance of the land cover, and beams this data back. The interpretation is done later and takes advantage of validated indicator relationships. Studies are now showing that there is great correspondence between land cover and "underlying" land use.

(2) <u>Self-Updating Data</u>

The above discussion leads into a fourth major technique for keeping data bases manageable: pinning data uses to data sources that have built-in mechanisms for updating. This does not mean that data themselves can rejuvenate. It does mean that even though data get old and out-of-date, whole new replacement sets become available from external sources. If the data base is structured with this in mind, file updating can be achieved almost automatically.

There are two broad kinds of coastal area data that can be easily kept current. They correspond to two of the primary information dimensions in the coastal zone. One of these is the land-cover dimension mentioned above. The other is a socio-economic dimension which can be thought of as "people-cover".

The earth satellite imagery can measure multi-spectral land cover to about one-acre resolution. This data, stored on digital tapes, can be processed by computer to cluster statistically-significant combinations of reflectances into "signatures". These can be correlated closely with land cover types through validation studies. Thereafter, validated signatures can be used to automate the interpretation of land cover.

The U.S. Geological Survey's Geography Program currently distinguishes three dozen major types of land cover for the United States. Correct interpretation at one-acre resolution is on the order of 90%. About two dozen of these land types apply in New Jersey: residential, commercial, industrial, transportation and utilities, industrial or commercial complexes, other built-up; cropland, orchard, other agricultural; deciduous forest, evergreen forest, mixed forest; streams and canals, lakes, reservoirs, bays and estuaries; forested wetland, nonforested wetland; beaches, bare exposed rock, quarries or gravel pits, transitional areas, and mixed bare land.

Depending on suitable atmospheric conditions, updating could take place as frequently as every nine days, now that two satellites are functioning. This degree of currency, and the low unit costs of the technique, could bring such tasks as the monitoring of land use change and general coastal zone surveillance within the possible domain of a coastal area information system (hereafter called a CAIS).

Another self-updating data source is the U.S. Census of Population and Housing. There has been a Census every ten years since 1790. There are also Censuses of Agriculture, Business and Governments which occur in intercensal years. Published Census tabulations are characterized by a wealth of detail and highest standards of accuracy and professionalism. The data can be used to form a comprehensive profile of the socio-economic characteristics of an area. Printed reports, though cheap and easy to use, do not include the full tabular breakdown of geographic detail. This detail is only available in the form of computer-readable magnetic tapes from the Census.

There are major advantages in having computer facilities at hand for Census data. A large portion of the coastal area of New Jersey (the counties of Middlesex, Monmouth, Ocean and Cape May) has no small-area coverage in printed reports. In addition, neither Ocean County nor Cape May had, at the time of the 1970 Census, been divided into Census tracts. From the 1980 Census on, summary tapes will contain small-area tabulations on all parts of New Jersey's coastal area. This means that socio-economic detail will be available at both the block-group/enumeration district and tract levels.

The major deficiency of Census data heretofore has been the ten-year interval between actual counts. The various techniques for estimating inter-censal changes have not proven wholly satisfactory in fast-growing areas. The current plans for five-year Censuses of

Population and Housing (from the 1980 Census on) would reduce the forecasting interval to something that can be reasonably bridged.

Five years is still a long time to wait for a data update. Fortunately, Census tabulation areas in New Jersey do not cross municipal boundaries. Consequently, any public agency that has an ongoing data gathering function and reports its findings by municipality can help to update the socio-economic profiles derived from census data. For example, the Department of Labor and Industry, the Department of Community Affairs, the Division of Taxation, and many other State offices regularly collect and tabulate data at the municipal level. Extensive machine-readable files are available for planning and modelling purposes, but are seldom used.

As it turns out, the variation in many socio-economic factors can be largely predicted in terms of two central variables: population and per capita money income. New techniques developed by the Bureau of the Census make it possible to derive municipal estimates of these variables with a high degree of confidence. Though unreliable in cases involving rapid growth or small geographic areas, the new methods provide municipal estimates at a standard of accuracy as yet unmatched by any other technique short of a full-scale count.

The thoroughly detailed base of data from the Census, plus annual and even monthly public agency updates, plus a reliable municipal inter-censal estimator are the basic ingredients for monitoring socio-economic changes.

(3) Data Structured by Geography

Effective computerization depends on the structure of data. Design errors in data structure can inhibit data use by forcing more time to be spent on format conversion than on productive analysis. Data format conversion is a chronic source of error and frustration; we believe it can be precluded.

There are many ways to organize data, whether by functional categories, age of data, source, scale, decreasing reliability, or alphabetical order. Each form of organization emphasizes different connections between factors -- maximizing some relationships while precluding or obscuring others. It is crucial to system usability that data be organized so as not to suppress known relationships, yet still enable innovative use. This means organizing data around a universal feature.

Land-based reference is one feature all data used in a CAFRA information system would have in common. CAFRA is not directed at unrelated events in a void. The whole purpose of the legislation, and therefore of the information system, is the joint relation of many factors operating simultaneously in one place.

To appreciate geographic data structure, it is necessary to understand the nature of geographic data. Every geographic datum combines a "fact" with a "location". The "fact" part may be thought of as a thematic descriptor, and the "location" part as a pair of cartesian coordinates. Technically, then, there are three dimensions to any geographical datum. The "x" dimension (east-west) and the "y" dimension (north-south) make up the location; the "z" dimension is a thematic descriptor (such as a count, measurement, density, or gradation). A thematic "map" is a display of some set of "z" values in terms of their "x" and "y" locators.

A data system for CA²RA could include any number of data sets, but would need only one geographic base. Whether this base covers only the coastal area or includes the entire State makes little difference, as long as all thematic data reference it in a uniform fashion.

The recognition of the common geographic dimension in CA²RA data uses simplifies not only the definition, but also the construction, of a Coastal Area Information System. Many different thematic files can share all or part of a common geographic base, thus minimizing the time and effort costs of bringing up new files for analytic use.

This brings us to the computerization of data. The three-coordinate conceptual format of land-based data allows each thematic "map" to be stored as a standard three-dimensional array in the computer. Referencing all thematic files at the outset from a common geographic base file routinizes data preparation, simplifies file maintenance, enables convenient data use, and avoids for all time the labyrinthine problems of data format conversion.

Considering the potential for a computerized data base, all Census data and much public agency data is already being collected in terms of geographic references. For example, the LUDA program of the US Geological Survey (see Appendix A-1) is now compiling satellite land cover data by municipality, Census tract, public tenure and major watershed. This means that a coarse environmental and socio-economic data base can be assembled and maintained at extremely low cost and high currency.

The foregoing discussion has sought to establish the logic, process, and feasibility of geographic data structure. The rest of this section attempts to show how this approach provides the foundation of a spatial information system.

(4) Pictures Worth a Thousand Words

Geographic data structure, besides being a logical and convenient way to computerize large files of land-based data, makes possible major advances in data productivity. Traditional computer processing too often merely transforms data into other data, without advancing the cause of meaningful information.

The methodological key is to take advantage of the interpretive power of the human eye. Geographic data structure makes it possible to utilize techniques of computer graphics to produce visual displays, or "maps". Compared to both textual and tabular presentation, visual displays have higher communicative efficency: they get the information across. They register meaning.

The human eye is a sophisticated pattern recognition device. The process of seeing involves the projection by the lens of the eye of a visual image on the retina at the back of the eye. In the space of just a few layers of cells, an undifferentiated image is composed into "pattern" and transmitted via the optic nerve to the brain which interprets the pattern. The brain can narrow the domain of recognition, focussing attention on small aspects of the pattern. This shift from the panoramic to the constrained field can be made at will. Reading is essentially the constraining of pattern recognition to a sequence of simple and specific shapes, relying on higher levels of mental processing to assemble meaning.

Reading, then, is linear processing. Simple shapes like the letters and numbers are compiled by the brain into meaning serially rather than as instantaneous patterns. There are many kinds of narrative, expository, descriptive, deductive, methodical, and logical communication techniques that rely on sequential processing to achieve meaning. It must be remembered, however, that all these techniques are really just a specialized and artificial constraint of native pattern recognition abilities.

Sequential data input to the brain is a relatively slow and inefficient way to register understanding of spatial interrelation. The more complex the terrain of data, the more difficult textual or tabular description become and the greater the value of the visual display.

The use of visual display is an important element in a land-based information methodology. By combining much factual and relational data into one piece of mental input, the essential "meaning" is preserved while the serial extent and total bulk of processing is reduced. In the theoretic terms of inherent information content, the use of visual displays combines data fidelity with compression. In the practical terms of an actual system, the use of visual displays means machines producing maps instead of people preparing text.

The geographer's goal of producing useful information without losing versimilitude applies to land-based information systems. To keep versimilitude high, land-based studies have resorted to increasingly refined data and complex models. Whether building better data bases or building better models for combining data, the result has been voluminous data processing and greater levels of precision, but no parallel increase in usable information. Computers were brought into land-based studies to handle the data processing and to maintain data accuracy, but have thus far contributed little to information methodology. The computer's facility for combining and transforming data into more data has made it all too easy to get trapped in a data avalanche.

Data fidelity is necessary but not sufficient in an information system. A strategy for data compression is also needed. Instead of the mere application of a computer <u>technology</u>, land-based information systems really call for the development of a computer <u>methodology</u>. It seems to us that the way to do this is to configure a system to bring data up to a point where the human mind can easily and rapidly draw intelligence from it, make meaningful conclusions, and generally reduce the data into information.

The appropriate device for data compaction and information distillation is not the computer at all, but the human eye. Graphics is the appropriate medium. Using computers to generate visual displays gives the end user far more leverage on information than does their use in standard tabular data processing. 1647 The superiority of graphic communication of spatial data is well known and not a new discovery. What is new is its potential for reorganizing land-based information systems. The reason for not relying more on graphics in the past was that manual techniques are expensive and time-consuming, and automated techniques were not yet available. Computer graphics have been slow in arriving. This is partly a market response, but it is also partly a consequence of the early digital bias of the computer industry. The heavy emphasis on mathematical serial processing capability in the digital computer was an artifact of the industry's early history. The early military and civilian applications of computers strongly influenced the subsequent growth and development of the industry. The late blooming of computer graphics is consequently more a historical accident than a logical consequence of information science. There is an incidental moral here: basic decisions made (or not made) early in developmental stages tend to wield great subsequent influence on future potentials for growth and expansion.

In any case, the computer industry has caught up in graphics. Whereas a few years ago technique was crude and technology was costly, recent investigation has shown us that sophistication and capability are high and heading higher while costs have dropped significantly. It turns out that there is no longer any good reason for dismissing computer graphics.

It may be that the historical orientation of land-based information systems to giant mathematical data processing capabilities was counter-productive. With the advent of mini-computers as remote terminals, it has become feasible to produce more useful information from less data processing capability. This is a good reason for paying attention to methodology early in the planning process of a coastal area information system (CAIS).

The keys to obtaining high information yield from a coastal area information system lie in using geographic techniques and computer graphics technology to take advantage of the human user's underutilized visual abilities. The following two section will elaborate on technique and technology.

(5) Computer-Augmented Analysis

The technique described in this section is a general way for making use of computer graphics technology and computerized geocoded data to display the indwelling "pattern" of spatial data. The major conclusion of our research is that end-product "information" and data display "pattern" can be made the same through user programming. The basic technique is presented in this section. It consists of a conceptual method with unlimited possible variations.

The common objective of the four DEP Coastal Area responsibilities is, as noted earlier, the simultaneous understanding of many factors in one space. A traditional analytic method for doing this is the "overlay map". This technique has been used by DEP, for example, in its Office of Environmental Analysis study of the Hoorter Quadrangle. This study consisted of the cumulative superimposition of data sets (overlaying of factor maps) to spot zones where specific sets of conditions occurred. The point of the study was to develop a method for identifying fragile ecological areas where performance standards should be applied.

The Hoorter study consisted of the manual collation of eight factors into a 7.5 minute USGS quadrangle sheet. Exclusive of first-time-only methodology and start-up costs, a single quadrangle of collated data cost about $15,000 in staff time and took four people a half year. The CAFRA territory comprises fifty-three such quadrangles. Based on this, a manual factor overlaying activity for the entire CAFRA area would cost $795,000 and take approximately one hundred person-years.

Fortunately, the use of overlays in geographic analysis is no longer restricted by tedious and time-consuming hand collation methods. Computers can be used to automate manual processes. Whole series of alternative overlays can now be generated in minutes, either on drafting paper in halftone shades for formal presentation, or on the equivalent of a television screen for interactive manipulation by an analyst.

The ability to overlay computerized data sets is an elemental capability of computer graphic routines. The real question is not what the machine does with factor sets, but what their superimposition means to the analyst who brought them together.

In most real-world situations of information need, no single data set, nor any single factor, has much meaning in its own right. Meaning is contextual. This is particularly true of meaning, or information, in a land-based context. To get geographic information, different kinds of geographic data must be combined. This is what a geographic information system does. This is also what geographic data displays do.

The process of stalking desired information is the business of the user of a geographic information system. In the hands of a good graphic analyst, information content can be made to converge with visual patterns of output data display. The basic methodological recommendation of this section is that general reliance be placed on the information-producing power of computerized geographic data overlays. We call these overlays "statistical planning surfaces".

The statistical planning surface is an output data graphic in three dimensions. Two dimensions handle geography, and the remaining dimension handles information content. The modulation of the third dimension can be handled in many ways (such as perspective altitude, half-tone grey scale, color hue, color saturation, patterned textures, or isopleth lines). As the choice of name shows, it is easiest to visualize the analytic output graphic as a three-dimensional map, or "surface". The statistical planning surface is essentially an automatically-generated "contour map" custom-made to yield information on a specific target question.

The "ridges" and "valleys" of a planning surface are not, of course, altitudes above sea level, but variations in some compound dimension relevant to land planning or management (for example, soil erodability, land value per acre, ecological fragility, or development suitability). The planning surface may combine any number of factor data sets as long as they are believed to relate to, "model", or "indicate" the spatial distribution of a land-based phenomenon being studied.

A geographic information system is based on the topological fact that perfect spatial coincidence in the data base will at least assure geographic correlation in any combination of data sets. The real skill of using such a system approximates model-making. Knowing what to select, how to define, and when to research are hallmarks of good technique. When the planner-analyst selects data for a planning surface, he is building a spatial model. The technique of using planning surfaces is virtually the same as that of planning: identifying the factors that are involved in complex geographic issues, and having a sense of the interrelation of data.

A question may have arisen as to how different data sets may be combined to produce a single display. This is not the forbidden act that it appears to be. The reason is partly that data are only combined with other data that refer to the same basic geographic location, and partly that data are not used in their own right. Data are used in graphic programs in a disembodied form. Actual units are stripped away and the resulting distribution of pure numbers is partitioned into scalar classes (using pre-set conventions of class interval and percentile distribution). Scalar classes can be readily combined and rules governing their combination made into an algorithm for computer processing.

The reason that data can be disembodied and combined as scalar classes is that they are all first judged to be "indicators" in some way of a common theme. The theory is that, though no one data set tells the entire story, several well-selected indicators can rough in a good spatial approximation. If data sets are truly relevant to an invisible target phenomenon, their joint spatial effect will tend to reinforce the "pattern" of location and distribution of the phenomenon being saught. The general effect is one of selectively reinforcing true information "signal" and canceling our spurious background "noise". Many types of variation are possible on the planning surface concept. For example, all factors in a spatial model can have equal weighting simply by being included. Or, the weighting of factors can be variable. It can be made to reflect known logical or mathematical relationships (an example might be using planning surfaces to predict locations of aquifer recharge areas). Weighting can also be used to build in pre-determined policy (an example might be special emphasis for certain soils to help preserve prime agricultural land from development).

The most valuable feature of the planning surface is that it preserves the basic informational content of data while reducing its digital bulk. Under even simple assumptions and coarse data, it is a powerful Its most interesting mode of application is factor combination according to decision logics (such as using only lowest or highest scalar values in each specific location, which makes it possible to evaluate land use in terms of limiting constraints and "holding capacity".

An imaginative user can adapt the planning surface to many applications. It can be used to identify and delineate zones (such as environmental sensitivity or housing need). It can optimize location or route choice (regardless of how "optimal" is defined). It can show a proposed land use change in its regional context (regardless of how regional context is construed). The planning surface may be a resource for merging major viewpoints or data profiles (such as socio-economic and environmental data).

Planning surfaces can be useful in the evaluation of alternative plans or program elements. They can be used to assess proposals for specific sites or find appropriate sites for specific land uses. They can also be adapted to illustrate the consequences resulting from series of alternative development scenarios. Because the weighting of factors may be changed, open evaluation and negotiation of trade-offs becomes feasible. It is also possible to make normative use of the planning surface concept to compute the gradient of spatial change necessary to arrive at a target future. This reverse use of planning surfaces makes it theoretically possible to do more efficient and comprehensive long-range planning and resource allocation.

All the above examples are feasible applications of a general information concept and technique. The versatility of the technique leads us to recommend it be made the central information methodology for DEP coastal area management. Though versatility and communication effectiveness are the prime reasons for a planning surfaces methodology, there are other important advantages too. They would include the following:

Understandability

Planning surfaces are as easy to understand as overlays. Because they are so readily understood, they have great potential application in the public information function. The conceptual simplicity does not detract from the utility of planning surfaces. Because they tend to capture the general equilibrium nature of ecology and economy better than text, the best role of planning surfaces is in policy formulation and comprehensive planning.

Economy

All the applications we have cited above simplify the process of doing geographic analysis. In the typical case, analysis consists of making straightforward interpretations of automatically-generated displays. The systematic use of planning surfaces in either planning or permit review analysis, could help routinize processes that have a tendency to move slowly. The integration of analytic graphics into daily routine could speed up analysis and free valuable staff time. The economy can be circular and self-perpetuating to the extent specialist talent is subsequently channeled back into the "design" or specification of further spatial models that can be used to routinize additional operations.

Continuity

There is little difference between creating an output display and generating a file for storage. It would be feasible to have some planning surfaces kept in a dynamic mode. It would be of value to permit review and long-range planning staffs in particular to have a continually updated picture of some critical land dimensions. An example would be maintaining a planning surface on the accumulated effect of incremental permit decisions. This would make it possible for staff to determine at what point effects projected for an individual applicant project would cause the projected cumulative impact of all projects to exceed established standards. The continuity of a dynamic model enables threshhold decision-making of this kind.

Consistency

General reliance on a computerized planning surface methodology offers a uniform way to analyze data coming into the planning or permit review process. It also offers a way to phase out old data, and to blend in new data. Rules and regulations backed by a rigorous and dispassionate analytic methodology can help build comprehensiveness into DEP processes and insure against charges of capriciousness or inconsistency. If, however, new data radically alter the contours of a planning surface, DEP would be consistent with the methodology to reject permit applications where similar permits had previously been approved.

Documentation

The planning surface graphic display is an effective kind of documentation. All that has been said about understandability and concise meaning apply just as much to documentary justification of a decision as to the process of reaching a decision. The spatial model that is used and the policy determinations of weighting, being perfectly explicit, present no problem. Just as pictures can be worth a thousand words, planning surface display graphics can save a thousand words. This could be of greatest value in justifying permit decisions and communicating policy determinations. Being able to show, for example, that a proposed land use change falls in the ten-percent of the coastal zone ecologically least suitable for development is a persuasive piece of evidence. Being able also to show as part of the public information process the ten-percent of the coastal area best suited to major development is persuasive information also.

Credibility

It will suffice to say that the universality of human pattern recognition, and the thoroughness and objectivity of computerized graphic display, give validity to planning surfaces and meaning to the patterns that emerge. The understandability of the method and the visibility of the results tend to give great credibility to the statements and conclusions that are derived.

We have explained man-machine graphic analysis, and recommended it as the appropriate answer to DEP's general information needs. We have earlier mentioned the superiority of this approach to either a wholly manual or wholly mechanical approach, and we have even briefly visualized a number of possible applications of the methodology. What remains to be discussed is the proper mix between man, machine, and process.

(6) Interacting with the System

In the foregoing section we explained man-machine graphic analysis, and recommended it as the appropriate answer to DEP's general information needs. We earlier mentioned the superiority of this approach to either a wholly manual or wholly mechanical approach, and we briefly visualized a number of possible applications of the methodology. It remains to make some observations about the optimal mix between man, machine, and process.

In our opinion, the more that technology can amplify the informational "message" by taking advantage of the human user's inherent capabilities, the better the information system. Up to now this report has been stressing the methodological advantages of bringing data into a computer-augmented format from which specific information can be extruded in terms of spatial pattern. Time also plays an important role in human perception.

Dual advances in technology and technique have dramatically shortened the time needed to produce computer graphics. Whereas Whereas traditional batch mode graphics with its slow queue turnaround can take hours or days, interactive graphics take seconds or minutes. Effective use of batch mode graphics depends precariously on ones ability to preserve a train of thought over long time spans that result from slow turnaround. This is no fit way to do analysis, and, fortunately, it is no longer necessary. Subsequent parts of this report, as well as Appendixes A-2, A-3, and F, will go into detail about equipment and technique. For the time, it will suffice to declare that present technology makes it possible to obtain a qualitative shift in the nature of an information system's utility.

Interactivity is a relative term. Three specific levels of technical capability seem to be involved:

- a relatively-inexperienced individual can use a computer system without the aid of an intervening specialist;

- this interaction takes place in the short time periods of the human attention span; and

- the system is programmed to be "conversational" in that the machine response cues the user into subsequent steps or options.

As system software is refined into something approaching language dialogue, as equipment is developed to respond with quick graphic output, as turnaround catches up with the pace of human inventiveness, and as the threshhold of necessary expertise is lowered to the narrow abilities of an infrequent user, an information system is doing more than just augmenting human capabilities. It is moving the user into a dimension that was never before possible. The system begins to be used as a direct extension of the users' thought processes.

IV. SPECIFICATION

The general information needs of the CAFRA functions form a "program" for a system. The preceding section presented the outstanding methodological expedients that we have uncovered. The present section concerns the process of turning the general program and methodology into a specific proposal.

(1) The Nature of the Alternatives

The range of alternatives open to DEP would appear to consist of doing nothing, systematizing only some of the CAFRA information requirements, or fully systematizing them all. Only the last of these turns out to be a viable choice. The first two are spurious alternatives. In the one case DEP can not afford to do nothing. Funding is already available, the Federal sponsor is expecting an information system, the permit review process is already operating near its full capacity, and the CAFRA timetable for planning environmental management is steadily elapsing.

In the case of the second spurious alternative, it is senseless to do a partial system when the threshhold cost for doing a full system is so low. A full system (one that can handle the similar information needs of of all the CAFRA functions) is the only real alternative.

We concentrated attention on systematic means for capturing and studying the underlying information patterns of geographic data. Spatial correlation is a more important fact of land-related data than is numerical behavior. Our approach to the program of a CAFRA information system has been in terms of information methodology instead of mere data processing. We recognize that this approach may be non-standard, but an institutional program as dedicated to geographic phenomena and spatial information as that of DEP is unusual too.

Inasmuch as there is so little real alternative to the full systematization of all CAFRA information needs, we believe a methodological approach is the correct one. Anyone who argues that CAFRA simply needs better data processing or just better data is missing the point. What is really essential is better information communication -- an analytic method that can make something meaningful out of otherwise meaningless data -- a flexible tool for condensing the "gestalt" out of voluminous data in readily perceivable ways that non-technical persons can quickly comprehend.

There is no alternative to comprehensive systematization if the program implicit in CAFRA is to be wholly fulfilled. From our investigation into methodology we conclude that this means achieving an interactive capability for manipulating, displaying, and analyzing geographic data. Anything less than this will not fulfill the program. The process of specification starts with this conclusion.

(2) The Nature of the Constraints

The most emphatic of the several initial constraints expressed by DEP was the injunction to avoid investigations that could be characterized as "state of the art." DEP's intention was an action program that would yield a usable system rather than a piece of original research. This turned out to be the best possible attitude because the state of the industry is such that all the constituent parts of a system are on the market, available in similar versions from multiple vendors, operational at one or more location, and readily implementable in New Jersey if so desired.

Other constraints we recognized were budget, personnel, and timetable. Of these, budget is the least constraining factor. Even though the anticipated level of federal funding has always appeared ample for unpinched CAFRA systematization, there is no reason why a complete system should be expensive. DEP already has access to the large and capable computer facilities of the New Jersey Department of Transportation. Compared to the basic DOT central processing unit and operating system, the additional software and equipment needed for CAFRA systematization is small. The dollar amount already budgeted in the Second Year Development Grant of the Coastal Zone Management Program for contractual acquisition of information system components is enough to do the job within one fiscal year. It is desirable to time major acquisition and development expenditures early in the federal funding to qualify DEP for subsequent administrative grants as rapidly as possible.

This is not to say that the budgetary process of establishing a CAFRA system would be either fast or unconstrained. The State Treasury imposes exacting requirements on bidding and acceptance. This insures that system purchases will be efficiently configured to achieve DEP objectives at least total cost. Detailed performance specification is necessary, but it takes time. Publication of the RFP takes time. Waiting for bids takes time. Choosing, or negotiating, an acceptable bid takes time. The principal consequence of the budgetary process is a substantial and unavoidable lag between deciding what is necessary and being able to get it.

This brings up the constraint of timetable. Other than for the mandatory budget process, CAFRA systematization could proceed at any speed. CAFRA only provides a general kind of schedule, but it is up to DEP to set the pace. Had DEP initiated the system acquisition process in January, a prototype system might already be operating and could have been used this summer to estimate and document the "capability of the various areas within the coastal area to absorb and react to man-made stresses." If the system became operational in 1976, it could be used to develop "alternate long-term environmental management strategies." If the system only became operational in 1977, it could at least be used to evaluate and select one environmental management strategy to be the "environmental design for the coastal area." In other words, the sooner a system could be implemented, the earlier it could be used in legislated CAFRA tasks.

It is useful to consider time more a resource than a dimension. In this regard, decisions must be made about how to invest time into studying, implementing, and utilizing an information system. Some time will be spent studying the needs and uses for a system. Some time will then have to be spent acquiring system elements and getting them together. Then time will have to be spent learning to utilize the system and bringing it into daily DEP operations.

Too often studies consume the biggest shares of both time and budget. DEP's broad program of activities calls more for general and flexible capability than for carefully-researched specialized routines for specified purposes. It should be possible for applications to be devised by users as needed from the capabilities of a versatile and easy-to-use system. We conclude that time should be invested in systems study only to the extent necessary to specify an appropriately versatile system and to justify the initial decision to build it. There is a point of diminishing returns whereafter further study just delays the time when actual use of the system might begin.

A coastal area information system (CAIS) will only be partly made of equipment and software. People make up the rest of the system. And people will be the most constraining factor of all. Systematization means automating expensive and time-consuming activities, removing redundancy, using present staff more effectively, precluding needs for future hiring, and minimizing the necessity of high-priced external consultants. Systematization also means innovation, or changing the ways people work. The reality of innovation is that it takes time to change things. Systematization is more than just acquiring new equipment and software. It is also <u>using</u> it in regular operations. Time is needed for staff to become familiar with new technique, to work out details of operation, to develop appropriate formats and procedures, to plan applications of technique, and to incorporate them into the institutional framework. The faster DEP can get a workable configuration of machinery and technique the sooner the slow task of system innovation can begin. Actual staff <u>use</u> of a system, even if of only a modularly-upgradable prototype, is the best way to invest limited time resources.

(3) <u>The Nature of Objectives</u>

The preceding paragraphs raised questions about further study and timetable that were not answered.

Regarding further study, if comprehensiveness is wanted, a computerized geographic information system is what is needed. If comprehensibility is wanted, interactive graphic analysis technique is what is needed. If both comprehensiveness and comprehensibility are wanted, the adoption of the methodology of this report is in order, and the steps for building and implementing a system should begin. No further study is necessary; a decision can be made. This is the course of action we recommend.

Though we would like this report to be as <u>persuasive</u> as we are <u>persuaded</u> about the desirability and necessity of such a system, we know that at first reading it is not. The major line of argument is going to seem theoretical and unconvincing to many. The only conclusive argument will be one that <u>shows</u> rather than tells how the proposed system can do better work than current methods.

Depending chiefly on DEP arranging staff time to do it, system visualization and prototype information graphics can be worked up concurrently with the publishing of a "request for proposal". In other words, steps to communicate the nature of a CAIS and to begin planning its applications could begin to take place in advance of having equipment and software physically installed.

Though it is possible for DEP to fully systematize its geographic information needs, it is not possible to acquire a "full" system externally. It should be remembered that the "system" is defined by its information objectives, not its physical parts. Equipment and software, which will come from outside DEP, are only part of the system. The other part, that which will actually determine system uses, is already "installed" at DEP in the form of personnel, data, and information needs. The decision to start implementing a geographic information system, then, means simultaneously laying out two things:

(1) the general target capabilities for equipment and software, and

(2) the applications by which users will get at complex indwelling geographic information.

As a way of beginning internal system development, this report, or an abstract of it, could be circulated to parties involved in the several CAPPA functions. A cover letter from the Coastal Area Planning Coordinator could request each concerned agency to propose from their work one typical planning or review problem that requires spatial determinations. These typical problems could be worked up as a series of quick-sketch case studies showing how the geographic information system proposed might automate and improve DEP techniques. This would serve four purposes:

- provide visualizations of the actual use of a CAIS in DEP operations,
- turn up potential strengths and weaknesses in the proposed CAIS,
- yield an initial checklist of typical problems a CAIS might be asked to handle, and
- possibly become a working outline for first system applications to information needs.

This brings up timetable--the second fundamental question relating to objectives. Is DEP behind, on, or ahead of "schedule" regarding an information system? The answer depends on how broad a definition of information requirements to accept.

Of the general mandates that delineate coastal area information needs cited earlier, the one that puts the greatest time pressure on DEP is CAFRA. The language of CAFRA does not, however, call for an information system by name.

It asks for an inventory in September 1975; for management strategies by September 1976; and for an environmental design by September 1977. If an information system is viewed as being just a mechanism for handling an inventory, time is more scarce than if an information system is seen as a major cog in management strategy or the main cam in environmental design.

The methodology (presented in this report) is well suited to sustaining all three levels of CAFRA implementation. Broadening the definition of CAIS to mean its being the vehicle for environmental design, not only gives DEP elbow room within CAFRA, but will also keep the "informing" role of information in perspective.

Despite our hortatory posture throughout the report, an information system really is a means not an end. Balancing ecology and economy in terms of land use is the goal, and a CAIS founded on an interactive analytic graphics methodology is a means to achieve the goal.

From this point of view, the report is proposing the methodological nucleus for a long-term environmental management strategy. It is a means for the "delineation of various areas appropriate for the development of residential and industrial facilities of various types, depending on the sensitivity and fragility of the adjacent environment to the existence of such facilities." It is moreover a way to "take into account the paramount need for preserving environmental values and the legitimate need for economic and residential growth within the coastal area."

One may conclude, then, that since environmental management strategies are not due to be fully developed until September 1976, DEP is ahead of schedule in considering the recommendations and methodology of this report.

V. TARGET CAPABILITIES

There are three general areas in which performance goals for a CAIS must be specified:

- inventory management
- quantitative research support
- interactive thematic mapping

The purpose of this section is to survey the major requirements in each of these areas.

Inventory Management

The creation and maintenance of a coastal area inventory involves problems of storage, editing and retrieval for two kinds of data: a geographic base file (GBF) to which all planning data are to be referenced; and comprehensive topical files which are the subject of planning research, e.g. soil types and economic profiles, well records and housing descriptions. It is anticipated that two different file structures will be called for, and consequently two different sets of file-management software. This dual nature of the data base in turn implies the existence of special software for file integration.

(1) The geographic base file

The GBF will be stored in the form of coordinate pairs (x,y) representing geographic locations. These pairs are identified either as isolated points, or as members of point-chains (line segments), or as intersections of chains. All coordinates will be given in terms of a common cartesian grid, most probably the NJ State Plane Coordinate System.

There are a number of ways of generating these data. The one most appropriate to CAIS needs is that of a minicomputer-controlled automatic digitizer (see Appendix A-2). This system of equipment and programs can for our purposes be treated as a single element, which as a whole must be able to:

- translate map inches into GBF coordinates
- rectify map errors in scale or orientation
- allow entry of labels (to identify segment or intersect)
- automatically close polygons, make "T" junctions, etc.

A somewhat more extended performance specification, prepared by the USGS for its own IUDA program, is worth examining in this regard.

The inventory based on the GBF produced by the digitizer must be accessible to logical search by its maintenance software. This means it should be possible to find all points within a certain distance of some line, or all bounded areas less than a certain size, or all lines that pass through a given polygon, or any other purely topological relationship. The only economical way to accomplish this objective is to have the spatial relationships built into the structure of the data base.

(2) Topical Planning Data

It must be emphasized that the GBF contains no "data" in the familiar sense of that term. It is only the "egg crate" in which the records of topical interest are stored. Or, to change metaphors, it is a bare Christmas tree on which otherwise uninteresting boxes of decorations are to be hung.

The other half of this dual-storage inventory system is the closet full of boxes that hold various items divided into topical (as distinct from spatial) categories. All the tax data in one file; all the soils data in another; census information in a third, etc. This file system also needs a package of managing software, typically referred to in the industry as "data base management systems" (DBMS).

The task of the DBMS is to facilitate the retrieval of stored data, the entry of new data or updating of old; and the inspection of logical and quantitative relationships among stored items. To this end, most DBMS packages consist largely of sophisticated routines for data retrieval and display, preferably with some built-in statistical programs for quantitative analysis. It should be possible for example, to produce a list of all municipalities whose school budgets were less than 80% of their budget, but which have total populations of less than 25,000 people, not including those where the median income is over $14,000.

Quantitative Research Support

It is necessary, but not sufficient, to be able to store and data at will. The CAIS must support planning and evaluation activities as well as simple inventory tasks. This means that its files must be not merely displayable, but manipulable--a vehicle for research. Most of this report is concerned with the use of controlled graphics as a primary research tool. But it is important to avoid the implication that CAIS should be limited to that approach. Graphic analysis is an enhancement, not a replacement for conventional tools of quantitative research.

(1) Statistics

A package of routines for exploratory and confirmatory statistical analysis is indispensable. The requirements of such packages are well established; the current market provides a spectrum of options, from "lightweight" conversational products such as STATPACK through "heavy-duty" multi-functional software systems such as SPSS (Statistical Package for the Social Sciences). The major variables that define the spectrum are:

- the size of the biggest manipulable file
- sophistication demanded of the operator
- complexity of available functions

SPSS, for example, foregoes conversational prompting in order to gain efficiency in the processing of very large files through multi-stage analyses (e.g. canonical correlation). At the other extreme, McNeil's IDA-APL can only accept data input at the terminal, but provides an unusual flexibility of operations, and assumes a wholly inexperienced operator. (See Appendix A-1)

(2) Modelling

The construction of mathematical models to reflect the interaction of many elements in a changing system is basic to three distinct but interrelated activities. Information modelling involves the application of previously tested equations to data of limited intrinsic interest, in an effort to expand their usefulness (e.g. dispersion models that extrapolate point-sensor data into area measures of air pollution levels). Such formulae can be altered or recalibrated by hypothetical modelling, in which data at point "B" are predicted by application of the model to data taken at point "A", the results being used as a basis for further research. Simulation modelling may be attempted with especially well-calibrated data-projectors, in an attempt to evaluate alternative proposals by "playing out" their potential results.

The system requirements for these activities vary from little more than a typical high-level language compiler (FORTRAN, ALGOL, PL/1) for information modelling, through the modification of special simulation packages (see, for example, summary of GPSS in Appendix A-1).

(3) General User Support

Most of the several applications packages discussed so far include formatting routines for the production of report materials in a manner legible by the non-specialist. Graphs, charts, histograms and statistical tables can all be produced automatically. But the explanatory text accompanying such exhibits must be prepared separately by the analyst involved. Of great benefit here would be one of the several text-management packages (e.g. ATMS/370) that enable the rapid reformatting of standard texts, as well as the correction and redrafting of material for subsequent presentations.

The activities dicussed so far--inventory maintenance, statistical analysis, quantitative modelling, and report preparation--imply more than one user being able to use the system simultaneously. This need may not be felt until the system has been wholly operational for some time, but it must be planned for. There are both equipment and software prerequisites for a multi-terminal system. In selecting an initial configuration for New Jersey's CAIS, the eventual implementation of these capabilities must be born in mind.

Interactive Analytic Mapping

The three-word title of this section spells out its primary elements. To provide an effective link among the various processes of planning, public information and project review, a CAIS must first be able to produce thematic maps. Second, it must be able to use these maps as a framework for policy analysis. Finally, analytic mapping must take place conversationally, in a format accessible to non-specialists, yet sophisticated enough to meet the demands of specialized investigators.

(1) Automatic Mapping

Two sorts of capability are required here. The system must be able to display points, lines, and polygons (PLP) with their associated "z" values--the latter either as numbers or in whatever graphic symbol the operator may deem appropriate (e.g. cross-hatching, colors, etc.) The system should also be able to display information stored as raster images (RI), with a similar range of graphic symbology.

It is especially important that the system have built-in utilities for switching back and forth between PLP and RI modes of storage and presentation. Specifically, it should be possible to extract the coordinate-pairs that define the boundary of any edge condition perceived in a rastered image; and to produce a rastered file representing the area enclosed by any string of coordinate pairs. PIP information is far more compact, and thus easier to store. On the other hand, BI information can be processed with special hardware far more efficiently than PIP data. CAIS should be able to claim the best of both worlds: processing in BI, storage in PIP modes.

In either mode, it is critical that the operator have control over the scale of map presentation, as well as its orientation (i.e. Mercer Co. need not fill the screen, and North need not be upward). Of equal importance is the ability to produce maps with appropriate labels, legends, borders and accompanying text necessary to assure easy and consistent interpretation.

There is a considerable choice of media for such maps. A standard line printer can be made to overstrike, giving a range of shades (as in SYMAP); a computer-driven pen plotter can produce line drawings of various colors and patterns, including cross-hatching (e.g. CALFORM); microfilm recorders can produce either vectored or raster maps both quickly and cheaply, and even in full color. But by far the most versatile medium for this sort of work is the electrostatic printer/plotter, a device distantly related to a xerox machine which can produce print or pictures under program control. (see Appendix A-4)

(2) Geographical Analysis

The simplest sort of geographic analysis is the displaying of some function of known data. If the display file is of municipal budgets, the shading might be adjusted to show percentiles, or budget as a function of relative geographic size. If several types of data are stored with reference to the same GBF (e.g. the local government survey), it should be possible to compute a relatively complex function (e.g. a ranking of municipalities by ratio of police budget to utilities expenditures, corrected for median income.) This function could then be grouped by deciles, or by distance from the arithmetic mean, or by distance from major cities, with the final groups displayed in several degrees of cross-hatching.

A somewhat more complex sort of analysis involves the generation of display areas other than those in which the data were gathered. The most familiar such technique is the contour or isopleth map. Here an assumption is made that the data vary more nearly continuously that the grain at which they were measured, and statistical techniques are used to "fill in" the data gaps. (CAIS should also have, as a subset of this general ability, a means for generating _proximal_ maps --a sort of contour which assumes a discontinuous variation of data, and attempts to project the likely location of the edge condition from point samples.)

The most valuable of the graphic analysis tools is also the most complex and difficult to achieve--the spatial overlay. The simplest use of overlay technique is to determine "nearness" relationships: what points of type "X" are within "N" units of line "AB"? These sorts of search function have already been specified for the GBF proper--it is important that they also be applicable to any image produced by CAIS out of that GBF. More complicated is the ability to perform logical comparisons between maps of the sort: what elements of map "A" correspond to given elements of map "B"? These could be the union or intersection of bounded areas, or the value of conditions at given points, or the parallelism of certain lines. The most sophisticated use of this technique allows the production of a complete third map "Z" which represents some function of map "Y" and some function of map "X" overlayed according to some third function, such as might be derived from a predictive model.

These three basic functions--the display of mapped information, the generation of new analytic boundaries and the overlay of one statistical surface on another--are the fundamental operations in geographic analysis, the axes of the system's capabilities. It is against the ideal of full and flexible use of all three capacities in both RI and PIP modes that any system proposal for a CAIS must be judged.

(3) Conversational Interaction

The analytic mapping capabilities described above have typically been used in a "batch mode" by the pioneer state systems (Maryland, Arizona, Louisiana--See Appendix C). This means that each step involved a trained programmer assembling a program in some such language as FORTRAN or COBOL or APL. If CAIS is not to be the restricted plaything of a handful of specialists, it will have to be "front-ended" by a powerful, interactive command language written especially for non-programmers.

A major barrier to quick acquisition of programming skills is the complex syntax of most programming languages. Conversational programming enables the construction of very complex commands represented by simple abbreviations--a twenty-line subroutine that shifts the center of a mapped display might be invoked with the single letter "C". This means that the operator need not know in detail the dozens of logical steps involved in creating an analytic map--he or she need only have a reasonably firm grasp on the issues to be investigated. The system can even be instructed to "prompt" the operator from the simplest of commands. Error messages can take the form of requests for clarification that guide one toward the proper format for analytic requests.

But this approach of preparing a long list of subroutines for easy operation is pointless if the operator must wait an hour or more for the computer to get to his task. This problem is much alleviated by such innovations as IBM's virtual storage operating system, which puts the system at the user's beck and call, and the new electrostatic print/plot technology that greatly accelerates the production of a monochrome graphic display. Truly <u>conversational</u> interaction with the system, however, implies another level of technological support, primarily in the mode of display.

What is needed is a display device with which an operator can interact directly. This is available in the form of television-type screens equipped with either a light pen or movable cursor with which the operator can indicate some part of the screen as that referred to in the present command. A still further development of this approach is the "menu" concept, in which the screen includes a table of symbols alongside the image, each of which refers to a system command. By pointing back and forth between the image and the command menu, an extended analysis can be performed by a person with no knowledge of typing, much less computer programming.

The extreme rapidity of this approach not only makes it feasible to use the computer to sharpen one's sense of the problem at hand, but also realistic for an analytic staff to explore dozens of alternatives prior to public discussion, culling the options that lead nowhere and grouping the others according to their degrees of similarity in impact, whatever their apparent difference in approach.

*

It is not imperative, it is not even reasonable, that all these performance goals be included in the initial RFP. A detailed and self-maintaining data base, an extended toolkit of methods and models, and conversational flexibility in all modes of geographic analysis will be the products of some years of successful system operations. But the <u>potential</u> to implement all these capabilities belongs in whatever preliminary configuration is used to create a CAIS for New Jersey. The choice of peripherals, applications software and support system must be guided by this image of what the CAIS is to become. No system is infinitely extensible. A performance horizon must be clearly visualized and planned for today, in order that it will still be achievable tomorrow.

VI. SYSTEM IMPLEMENTATION

Thus far, we have been concerned to make three main points:

- The CAFRA/CZMA mandate to provide new functional links among planning, review, project development, and public participation implies an information system capable of integrating a wide range of users, data, and problems.

- Interactive geographic analysis provides a "common tongue" in which that integration might be achieved.

- A Coastal Area Information System based on this approach (but also capable of the full range of library and calculator services conventionally associated with "data processing") could be assembled from elements commercially available in June 1975.

We must now assess the conditions under which an operational CAIS could be implemented, given the familiar constraints of time, money and manpower.

The CZMP Second Year Work Plan calls for an operational CAIS by April 1976. (This is consistent with the CAFRA timetable, which calls for the presentation of alternative environmental management strategies--presumably based on CAIS-supported research-five months later). Counting back from that goal, we must allot:

- 30 days after delivery for equipment shakedown

- 60 days after contract award for equipment delivery

- 10 weeks to evaluate bids prior to contract award, and

- a minimum of six weeks for circulation of the RFP

To the extent that this constitutes a critical path schedule, work should start immediately on CAIS performance specification, and a request for bids should be initiated in late October or early November, for the sequence of system implementation to conform with the proposed CAFRA and CZMA work plans.

These points provide the outline for a streamlined implementation process. A list of the major work elements or activities included in that process is contained in Appendix E. That list is followed by a nine-point sequence of conditions that would have to be met if CAIS implementation were to be accomplished within the mandated time-frame. The following comments are arranged according to that sequence. They are thus somewhat in the nature of a trail guide: notes on the dangers to be expected along the way, including some tips or how some of the more troublesome quagmires might be crossed--or avoided.

(1) Accepting Conclusions of this Report

If the three points listed above cannot be accepted, then we have nothing more to say. Other consultants may propose other sytems. IBM has suggested 5 calendar months as a minimal time for a system study--a figure we find reasonable. If some strategy other than that which we are proposing is to be chosen, then the appropriate study should begin as soon as possible. That would probably make January 1976 the earliest that procurement paperwork might even begin.

This brings up the second hurdle in this category. The Bureau of Data Processing Management (Treasury) is charged by law with the supervision of all procurements involving computer equipment and/or services. If this office is opposed, either in principle or on any large number of specifics, to the strategy being advocated here, all CAIS-related activities might get postponed until an acceptable compromise is reached. The vital thing to keep in mind is that technical considerations and negotiations not inadvertantly undermine the objective of a versatile geographic analysis capability.

(2) Creation of Staff Nucleus

A system implementation staff can either be assembled by reassignment of current staff, or else hired in newly created positions. The latter takes more time. Three sets of skills should be distinguished as necessary to the staff:

- systems specialists for the data processing gear;
- planners and other departmental analysts;
- a moderate level of technical/clerical support

It is not critical that all of these be full-time members of a CAIS task-group, but continuity will suffer if there is not some representation in each of these categories within the permanent CAIS staff.

Crosscutting these three groups is the critical distinction between development and operations. The functional importance of maintaining these as discrete tasks is discussed in (9) below--it is only noted here that there is a parallel range of personal styles and competences:

- Technical imagination, a generalist background, and unflagging optimism are the keys to fertility in developmental work--yet they can spell disaster in an operational environment;

- Conversely, only methodological rigor and deep specialization, along with a cynical assumption that even what can't go wrong probably already has, can sustain daily operations--an approach that hamstrings attempts to extend and improve system capabilities.

Hiring and/or assignment must be done with an eye to matching individual proficiency with the right class of task.

Finally, it would be unwise to place the management of so complex an effort in the hands of anyone unable to give it constant attention. The same could be said of any attempt to locate the CAIS wholly within some extant office with prior (and inevitably competing) commitments. The argument applies _a fortiori_ to the possible distribution of CAIS subsystems among several such offices. No course can be wholly secure, but the management of CAIS development and implementation would seem to be most hopefully assigned to a full-time professional who reports directly to the CAFRA or CZMA Program Manager.

(3) Sequence and Priority of Objectives

Timing and attention are closely related. As a rule, the more important tasks should be addressed first; but the step-by-step necessities of system development are also known to determine the relative importance of different tasks at different times.

As soon as the funds are committed and the staff chosen, it will be critically important to establish a firm image of how an operational CAIS will serve the multiple mandate of planning and review, project design, and public involvement. The concrete objective here is a priority list of target capabilities to which calendar dates for implementation have been assigned: how soon will each service be expected, and how is it to mesh with ongoing activities? Clear answers to these questions will go far toward smoothing the transition to automated operations.

(4) Preparation of Bidding Specifications

The number of imponderables associated with major State procurement is legendary. It is at this stage where the consensus discussed in (1) above becomes critical--and where the timetable for system implementation is most vulnerable. The following points are offered for consideration:

- There may be no single vendor willing to bid on the whole system, which could greatly complicate both system integration and all aspects of maintenance.

- Many vendors are impatient with the strict acceptance criteria that typically accompany State contracts. Unfortunately, these include some of those most experienced in geographic information systems--raising the possibility that the most qualified vendors may offer no bid.

- An overly restrictive RFP runs the risk of merely attracting a congeries of partial bids from unknown vendors, overlapping some portions of the system, and missing other portions altogether.

- Because this whole area is based on the application of already standard software to new sorts of problem, it is experiencing an explosive growth rate: since our March presentation, we have learned of three new developments by major vendors--each pointing in a different direction. This also suggests that the specifications in the RFP should be written generally enough to attract a wide range of vendor replies.

(5) <u>Evaluation of Bids and Award of Contract</u>

An RFP such as we propose will produce bids singularly difficult to evaluate in simple cost/benefit terms. The following should be kept in mind:

- the functional core of the system is its ability to interface graphic and file-management operations: specifically, to overlay and contour mapped data, and to generate new files from the images that result.

- all other aspects of system equipment and programming will thus be determined by the requirements of the software that integrates the digitizing, data base management, and graphic display modes of operation.

- There are at least two fundamentally different approaches to achieving this software "spine": one is built around the description of points, lines, and polygons (PLP); the other is derived from raster-oriented image enhancement technology (RIE).

- The RIE and PLP approaches to interactive graphic analysis ultimately provide the same functional capabilities, but their starting points are quite different. RIE graphics begin with a high-resolution image (stored and manipulated in special hardware), from which the geographic coordinates of particular features are extracted; PLP graphics begin with a file of geographic coordinates from which an image is built up one line at a time. Consequently, the "primary configurations" under each approach are dissimilar.

- A fully implemented CAIS should have both RIE and PLP capabilities. But the choice of which should be used as a system starting-point should be postponed until bids have been received. Contract award should be based primarily on two considerations:

  - demonstration of efficient and flexible interface between image and file structures--in particular, the overlay and contour capabilities; and

  - evidence of system extensibility to include both PLP and RIE modes of operation.

It should also be noted generally that device-independence, a traditional goal in software development, is an inappropriate criterion for evaluating proposals in this field. Virtually all of the work that has been done has involved the design of special equipment, which is in turn driven by unique, hardware-specific programs. (This is conspicuously true of RIE technology). Emphasis should therefore be placed on the vendor's record of preserving system compatibility across changes in hardware design, rather than on evidence of software transferability across vendor lines.

(6) Data Collection and Organization

The shape of this task will depend on the relative emphasis given to PLP and RIE data during the first year of CAIS implementation, and on the priority-sequence of system objectives. The distinction between two phases of the task will, however, remain constant: deciding first what to collect, and then compiling the chosen data into a uniform set of files.

- Data Selection: criteria for this phase have been discussed at some length in section II above, and are summarized in Appendix B. They include geographic scope and resolution, file reliability and extensibility, data relevance and expense.

- Data Organization: again, a twofold task, involving (first) a determination of the logical relationships among file types and subject matter--a task which should lean heavily on the Coastal Zone Information Source Profiling System (CZISPS) and the Coastal Zone Environmental Inventory Subject Index, as drafted by the DEP/OEA--and (second) formating the files according to system specifications, once the contract has been awarded.

Six months is a plausible estimate of the time required to complete this task. It is desirable to assign a task force to begin work in this area as soon as the RFP has been published, or perhaps as early as October.

(7) First System Application

Consistent with the bias toward planning evident throughout this study, our recommendation is that high priority be given to the delineation of resource-management zones, expanding on the example set by OEA's experimental study of the Boonton Quadrangle. These could then become the basis for a set of management criteria specific to both resource area and type of development--providing a guide for permit review as well as a basis for public comment.

The zones and their associated development criteria would be provisionally accepted, and used to structure the first system-oriented review of permit applications, as well as the first demonstration of the system's accessibility to public use and evaluation. One advantage to this approach is that the first products would be something at once concrete and modifiable. Another is that the basic conceptual work has already been done, thus minimizing the uncertainty associated with the first implementation of a new analytic method.

The design, preparation and first-stage implementation of such a project would occupy 2-3 staff members for perhaps six months. This implies a third specially assigned task-force within the CAIS development staff (aside from those working on bid evaluation and data organization), also to be created sometime in September.

(8) Equipment Installation and Shakedown

Our industry survey indicated equipment delivery can be anticipated sooner than the quoted 60 days after contract award. The vendor can be expected to uncrate and install their own equipment, and perform a basic battery of diagnostics to show that everything is in working order. But that is still a long way from having an operational information system.

The most difficult aspects of implementation come under the heading "system integration." Data, software and equipment must be brought together amid complexities that can be generally predicted but never adequately specified until they are confronted. Our investigation suggests that a good month should be allowed for sorting out these difficulties and getting the most quarrelsome of them enough under control to permit some semblance of normal operations. The implementation of the data base (or at least some significant portions of it) is included in this estimate, but not the debugging of any new applications software.

(9) Operations and Ongoing System Development

The final item on our list of implementation quagmires has already been alluded to in (2) above. It involves an issue of enormous urgency, which is nonetheless routinely ignored in State agencies: the distinction between operations and development.

Lip service is generally paid to the importance of developmental efforts, but transforming words into action is another matter. Developmental activity is frequently the victim of administrative shortsightedness. On the one hand, manpower may be cannibalized for projects of alleged immediate necessity. On the other hand, development efforts may languish for want of adequate contact with the flow of daily operations. The proper function of a development staff is not fighting brushfires, but devising better tools for meeting--and preventing--future emergencies.

There is a price for treating development staff as a reserve that can be routinely drawn upon to meet operational pinches. It is paid by the continued recurrence of pinches that might have been permanently prevented by successful development work. A still higher price, however, is that of program stagnation. A program stalled at a fraction of its capacity both wastes present reources and squanders future capability. This is nonetheless a far from uncommon occurence. Early success leads to an increased operational load, which in turn leads to more frequent bottlenecks and production pinches. If these are met with repeated drafting of development staff, the vicious circle is complete. A development staff cannot achieve what it is never given the opportunity to work on.

It is helpful to distinguish three sorts of staff effort: operations, development, and emergencies. Under normal circumstances, staff energies might be devided 60/40 between operations and development. During emergencies, three quarters of the development staff could be drafted to cope with the situation--but no more (if a 90% staff effort cannot handle the problem, there are probably too many people getting in each other's way). Afterward, however, this deficit must be repaid in the form of operations staff made available to recoup the developmental momentum lost during the emergency period. Conversely, when the time comes to implement a system refinement or extension, an operational backlog may be deliberately created, as operations staff are diverted to manage the implementation.

The tension between efforts to get today's job done and efforts to make next month's or next year's tasks better and easier is both fruitful and inescapable. Any attempt to remove that tension will be detrimental to project success, sooner or later. The goal must be to create a working environment in which both tasks can be vigorously pursued, and where their paths must repeatedly and knowledgeably cross. Beyond this, there is no purely administrative solution. The success of a project ultimately depends on the caliber of its staff.

VII. CONFIGURATION

The three foregoing sections discussed constraints, target capabilities, and implementation of an information system. These sections constitute a springboard for proposing a configuration for a system. The question at hand is what constellation of equipment, technique, and personnel will achieve the desired capability rapidly enough to be useful in CAFRA program development.

Our analysis concluded that the dimension of system development most difficult to control was time. There is greater certainty about what industry can deliver than what staff will be able to accomplish. That is why it is desirable to arrange total system development in such a way that as much time as possible is provided for staff uses of the system. It should be kept in mind that the major initial use of a CAFRA information system will be in developing DEP human resources. An information system does not offer a substitute for skillful human analysis. It only eliminates redundant operations and expedites slow processes. The machine's role in the information system is to augment human processes and amplify human skills. A geographic information system is fundamentally an environment for man-machine interaction with land-based data. That is why consideration of system <u>configuration</u> can only come after fundamental decisions have been made regarding the capabilities and methodology of eventual system <u>use</u>.

The primary thrust of this report has been to present a high-yield methodology for producing geographic information from raw land-based data. The second thrust of the report is now to reiterate a previous finding that a geographic information system can be made operational at DEP in the coming months in about the same amount of time and money as a system study might have taken.

On March 18, 1975, the Division of State and Regional Planning proposed a feasible configuration for an information system that could be implemented quickly, could reasonably support DEP information needs, and would represent the state of the industry. Our conclusion was that this information system did not have to be expensive. In fact, the major hardware item was a <u>d</u>edicated "mini-computer" <u>s</u>upport <u>p</u>rocessor (DSP) which could be purchased outright for as little as $43,000. This is not to say that our proposal found a way of avoiding the need for the large data processing capability of a big computer <u>c</u>entral <u>p</u>rocessing <u>u</u>nit (CPU), but it is possible for a DSP to be connected via a teleprocessing link to a large CPU. The DSP would rely on the CPU for complex processing, and handle simpler processing by itself (thereby not degrading the operation of the CPU).

Our proposal recommended that DEP could take advantage of the good CPU facilities already at the Department of Transportation Computer Center and get along with just a DSP and associated pieces of peripheral equipment. This approach is far less costly than purchasing a self-contained turnkey system, and also avoids unnecessary proliferation of State computer facilities.

Appendix F contains the textual material from the March 18 presentation. The diagram with which it begins is a useful visual aid and should be consulted at this point. It is a schematic representation of the DSP (central box), software (internal rectangles), the DOT computer (upper lefthand corner), and remote peripheral equipment (circles). We suggested that this constellation of parts could support all target capabilities and probable CAFRA uses. We called this configuration a "feasible" one because the various elements were available on the market, could be gotten in most cases from several manufacturers, and were being used successfully at other installations. used at other installations. Employing generous estimates, we stated that all elements of a system could be acquired for approximately $250,000. And if phased implementation necessarily had to be chosen, this cost could be spread out considerably.

Since March we have received supplementary information that has confirmed us in the conviction that a sophisticated information system can be assembled for a relatively modest price. In fact, one manufacturer has announced a less expensive magnetic disk and a better and more efficient operating system. This leads us to believe that the March configuration could now be acquired and implemented for as little as $200,000. Another development is the recent availability of in-house image processing at low cost. This means that for about the same $250,000 cost as we originally suggested, a great deal more interactive versatility and "analytic power" could be achieved.

The textual part of the March 18 presentation contained in Appendix F is difficult to understand divorced from the verbal explanation that accompanied it. We have consequently provided some of the technical vocabulary and background for understanding the configuration and the diagram in Appendix A-2. For the purposes of this concluding section, it is preferable to characterize the system configuration in only the broadest terms. There are three major dimensions to the system. It is land-based; it is graphic; and it is interactive. Each of these dimensions requires a subsystem of equipment and software to sustain it.

The land-based subsystem consists of the software structure that organizes data, and the equipment that makes it possible to bring new data into the system. A major part of the land-based software support system is the geographic base file (GBF). The GBF makes it possible to relate all data in terms of geographic (x,y) identifiers, and to manipulate primary data elements conveniently and efficiently. Equipment and software are necessary for converting mapped and tabular data into a machine-readable geo-coded data format. The same capability serves for mass entry of new geography as well as for the selective alteration of the GBF. It is called the digitizing and editing capability, and can be run on a mini-computer.

The graphic subsystem is the equipment and software that enables users to display data in different ways in order to make the spatial distribution evident. Graphic display, by taking advantage of natural pattern recognition abilities in the human eye, is a giant step toward automating the analysis of land-based phenomena. It is the easiest way to cross the gap from data to information. The most cost-effective hard copy graphic display device is probably the electrostatic printer-plotter, and it can be operated on either a mini-computer or a remote extension of a larger computer facility.

The interactive subsystem is principally a question of software enabled by appropriate hardware. A rapid volley of instructions and responses makes it possible for the user quickly to approach an understanding of data. This is especially the case if interactivity can be achieved in conjunction with graphic display. In the case of interactive graphics, two terminals are needed: one for output displays, and another for transmitting to and from the main computer. Interaction can be supported on a dedicated mini-computer or a remote time-shared system.

All three subsystems can be dimensions of a single operating system utilizing a dedicated mini-computer with a remote job entry (RJE) teleprocessing link to the big DOT computer. All CAIS operations would rely to some extent on this main CPU and DOT's Conversational Monitoring System (CMS). The extent of reliance would vary from digitizing (where the reliance would be minimal) to modelling and statistical programs (where processing would take place totally at DOT with only output coming back).

This kind of configuration avoids the major cost of having to set up a whole new computer system. It takes full advantage of the large and efficient DOT computer facility (an IBM 370/145 VM machine), and thereby dramatically reduces the cost of achieving the CAIS target capabilities. The only hardware expenditures would be the DSP, the teleprocessing link to DOT, and the peripheral pieces of equipment (data storage units, communications terminals, digitizing table, and "hard" and "soft" graphics devices) indicated in circles on the diagram in Appendix ?.

This brings us to consider further specificity. The state of the industry is in a constant state of flux. The relative position of one vendor vis-a-vis another vendor is impossible to fix. New software and hardware products are continually being introduced,new capabilities of current products are being developed, new vendors are entering the market, and the price schedules of current vendors change. There is no point in trying to achieve much further specificity in this proposal until after the fundamental commitment is made to build a CAIS. Further specificity, when it does take place, should take the form of detailed performance specifications worked up in conjunction with the professional systems staffs of the New Jersey Departments of Treasury and Transportation. The ideal case would be to develop and publish only one RFP. This Request For Proposal would cover all target capabilities of a CAFRA information system, and would make potential system elements the responsibility of the bidding firms.

Soliciting competitive bids on a single system "package" is an effective way of substituting capital for time and labor. A single package would cost more than the sum of the individual elements. But what the increased price buys is somebody else's time solving the technical complexities of bringing the various elements together. There are some headaches worth avoiding, and some prices well worth paying. Competitive bidding on a carefully specified RFP is probably also the fastest possible way of generating a series of detailed variations on configuration. This is more than any single "study" can hope to accomplish, and the practical path we would urge DEP to take.

This brings us back to the objectives of this study. It set out to draw attention to the recurring needs for comprehensive data, analysis, display, and documentation that characterize potential planning, permit, and public information activities at DEP. It pointed out overlapping information needs of the public, the planning staff, and the permit review staff which constitute a program for a system. The report proposed DEP adopt a comprehensive solution to this program in the form of a geographic information system based on interactive graphic analysis. It also suggested basic target capabilities for a system and the pitfalls and priorities of implementation. A feasible configuration for a CAIS was generated that can reasonably be expected to satisfy DEP information needs as well as provide a versatile medium of analysis and communication.

This report has tried to show that building an information system is more than just automation and data processing. A comprehensive solution to the coastal area information program involves a configuration of staff, equipment, technique, data, and methodology. A finding of this report is that the resources, technique, talent, and equipment are available if DEP can provide the executive backing and staff to carry forward a unified plan. Though there is not one topic in this report that was really covered thoroughly, enough has been said to warrant some fundamental determinations. A small number of key decisions made now can set DEP on the path to realizing a unified information system for coastal zone management.

APPENDIX A:

THE STATE OF THE INDUSTRY

Our approach to specifying the design of a CAIS has been based on a postulate that the hardest problems in geographic analysis have already been solved: that "somewhere out there" the necessary equipment, software, and data-handling techniques have been developed and are already in use. After extensive research into this subject, our convictions remain unshaken. The purpose of this appendix is to convey our perceptions of the state of the industry, and to document our assertion that there is more than enough technology and technique readily available in the market to provide a comprehensive geographic information system able to satisfy DEB's basic needs for geographic information.

Our work began with the identification of available systems and packages specifically designed for automated geographic analysis. From these, it branched into a consideration of other general-purpose programs that by convenience or necessity are used in conjunction with those specialized products. Our findings in this area are synopsized in Section 1 below.

That survey of systems and packages inevitably led to a study of the machine requirements associated with various system options. This eventually involved a broad survey of specialized equipment for automated cartography and interactive spatial analysis. The results of this investigation are outlined in Section 2.

Considering the great methodological advantage that attaches to graphic analysis and display techniques, examples of output are included and make up Section 3 of this appendix. These examples should only be taken as representatives of a broader class. The important thing to keep in mind is not what these examples specifically communicate, but the general capability that they embody.

Section 4, the final part of this appendix, summarizes the major implications on CAIS specification we have drawn from studying the state of the industry. In our opinion, there is not a simple continuum of capabilities and expenditures to be considered. Each additional dollar of expenditure does not secure a slight bit more capability. Rather, the state of the available tools and the requirements of system integration limit the meaningful options to three plateaus of configuration. Each plateau represents a magnitude increase in commitment, and each implies a magnitude increase in the present and future role for a CAIS.

APPENDIX A: Section 1

Systems and Packages

This section is a list of pointers, not an exhaustive description. There is much that could be appended to the following catalogue, and much more that might be said about each item listed. Further information is available to DEP on each item listed: in our DSRP files and also from the contact person indicated after each product name. We have consequently stripped our findings to a non-redundant characterization of the kinds of capabilities available on the market. Were DEP merely to implement some few of the following packages, and not to devise a single addition or improvement to software or equipment, DEP could still go a long way toward realizing its most ambitious information goals.

The following is the list of representative systems and packages selected to illustrate the existence of versatile display and analysis capability. In the second column is a more useful list--the acronyms by which these systems are commonly called:

. A Design and Planning Tool	(ADAPT)
. A Programming Language	(APL)
. Cartographic Digitizing/Editing System-PDP8F	(CART/8)
. Geographic Data System	(GEODATS)
. General Purpose Simulation System	(GPSS)
. Harvard Laboratory for Computer Graphics and Spatial Analysis:	(LCGSA)
. Synagraphic Mapping Program	(SYMAP)
. "Calcomp" Plotter Conformant Mapping Program	(CALFORM)
. Synagraphic Mapping 3-D View	(SYMVU)
. Polygon Converter	(POLYVRT)
. Interactive Exploratory Data Analysis	(IEDA)
. Land Satellite (successor to ERTS)	(LANDSAT)
. Earth Resources Technology Satellite	(ERTS)
. Land Use Data Analysis	(LUDA)
. Maryland Automated Geographic Information System	(MAGI)
. Multi-spectral Data Analysis System	(M-DAS)
. Statistical Package for the Social Sciences	(SPSS)

ADAPT (Aerial Design and Planning Tool):

A package of FORTRAN IV routines developed by W. E. Gates and Associates, Inc., of Cincinnati. Written to assist in water quality management, the package a large number of quantitative models for predicting water pollution from land uses. ADAPT has an unusual triangle-based GBF, but interfaces the Harvard LCGSA mapping packages for spatial display. Originally, it was a batch-mode system that ran on an IBM 370 computer. The Gates company, its author, is now in process of implementing ADAPT to run off a mini-computer. The client is the State of Kentucky. In both the large-CPU and minicomputer forms, ADAPT will be available in the public domain. Its major weakness is a crude digitizing/editing capability.

APL (A Programming Language):

A highly sophistocated computer programming language that was designed for interactive problem analysis. It is very suitable to the kind of complex, exploratory data analysis that is entailed in a geographic information system. One of the most important features of APL is its great concision which gives user straightforward and powerful information capability without the byzantine format and syntax problems that complicate many other programming languages (such as FORTRAN). APL's odd character set and extreme condensation may frustrate dilettante users, but in regular usage these characteristics turn out to be APL's most attractive qualities.

In the Trenton area alone, there are at least three different kinds of APL applications programming in operation. They show the general range of APL's versatility:

- FPLOT is an IBM-sponsored product capable of producing curvilinear graphs using a special "selectric" type element on a standard typewriter terminal. Graphics have a resolution of 0.1 inch.

- TMODEL is a set of interactive graphic routines for modelling transportation systems. It currently uses Trenton as its test case. It is frequently used in real-time public presentations and is a good example of a public participating application.

- IEDA-APL is an interactive package for "exploring" data. It was developed at the Princeton University Statistics Department, and is a major contribution to graphic analysis. It basically is a way to display principal statistics in a visual shorthand. By facilitating easy perception of intrinsic data behavior, the user can quickly advance to statistically-superior expressions of data that are both true to the information content and useful in modelling.

APL is available both at the Princeton University Computer Center and the Educational Computer Network of the State colleges. Mr. Barry DuSault (201-932-8051) is the knowledgeable expert and a congenial exponent on APL.

CART-8 (Cartographic Digitizing/Editing System):

An extremely refined system for interactively generating and maintaining GBFs (Geographic Base Files). It was developed at the University of Saskatchewan and is marketed by DYNAMAP, Ltd. The staff of the LUDA program at USGS highly praise CART-8. The "8" of the acronym refers to the PDP-8E minicomputer that runs the system. The hardware consists of a digitizing table of the free-cursor variety, a graphics terminal equipped for complete operator interaction, and a dedicated minicomputer.

The operator can digitize data from a map or photograph, and the lines appear simultaneously on the screen of the terminal for inspection. Corners and "T"s are corrected by the system software, which can also store labels, coordinates, and scaling information, as well as modulate the digitized information to conform to any map scale in any orientation.

The system is expensive (about $130,000 fully equipped), but also remarkably thorough. Because it is designed to be a stand-alone device (see notes on other digitizing systems in Appendix A-2), CART-8 is unsuited for integration into a multi-funtion CAIS in its present configuration.

Good people to contact about CART-8 are Richard Prytula, DYNAMAP Ltd., Saskatoon, Saskatchewan; and Robin Fageas or Bill Mitchell at the U.S. Geological Survey in Reston, Virginia.

GEODATS (Geographic Data System):

Is the most sophisticated geographic information system we have encountered. It can accomplish the largest number of functions with the smallest investment of equipment and programming. It includes all of the PLP capacities discussed in this report, but does not yet include any RIE capabilities (see Section of report).

GEODATS is marketed by a new group called ESCA-tech (formerly GEO-science) and is installed in booming Alaska at the Chalisto Native Corporation. ESCA-tech is experienced in image processing and photo-interpretation of aerial-photography. They are in process of acquiring raster image capability.

From our vantage point, they are the standard against which other systems should be measured. Although GEODATS is advertised as a completed package, it is understood that each installation of the package is unique. Special interfaces are required for each computer. File design is to some extent dependent on the nature of the files to be processed. Moreover, each client will have priorities among the various functions it requires. For example, ESCA-Tech has not yet had a client with an interest in contouring as a part of a geographic information system. ESCA-Tech does have a contouring package developed for other purposes which they claim would be easy to interface into GEODATS for New Jersey use. Other possibilities include interfacing to already available data bases such as LUDA (See Below) for the use of LANDSAT imagery. The cost of GEODATS would in large measure depend on the specifications of the package: what functions would be done locally (GEODATS is designed as a stand-alone mini-computer system but can be a remote job entry into larger systems), what functions would be performed either by a remote system or by a service Bureau, what data files would have to be accommodated, and what equipment would be purchased from ESCA-Tech as part of the system contract. Published numbers range from $80,000 for a listing of all of the software through $400,000 for a system design from scratch. New Jersey's needs fall somewhere in between.

GPSS (General Purpose Simulation System, release 5):

We believe the future of computers in the support of public policy design will not rest primarily on their abilities as super-librarians or super-calculators. Rather, it is the possibility of simulating the interaction of many variables under many circumstances that will make computers a powerful tool at the service of decision makers. Most models designed to simulate a particular set of events are developed on a case by case basis. It is, however, enormously helpful to have available a language geared especially for creating such models. GPSS is such a language. It is also probably the best documented and most widely used simulation package. A number of books have been written on the use of GPSS in operations research where the object of management may range from a river basin to a factory. The value of packages such as GPSS is that the operator or programmer is free to concentrate on the nature of the modeling problem rather than on the details of programming code. In this sense GPSS is similar to the many statistical packages available for computer use. Instead of each GPSS command invoking a routine for statistical analysis, it calls a routine relating events to conditions in particular ways. A similar, though more sophisticated product, is SIMPL/1. The Trenton IBM office is the place to contact for information (George Cummings is particularly helpful and knowledgeable). GPSS(V) and SIMPL/1 are available at the DOT and ECN computer facilities.

LCGSA Products (Harvard Laboratory for Computer Graphics and Spacial Analysis):

(1) SYMAP

This is the oldest and best known computer mapping package-SYMAP uses as its output device the line printer of any computer. It creates its shadings by overstriking characters to create a perceived range of gray shades. Although its output appears crude by the most recent graphics standards (such as microfilm), its information handling capabilities are unexcelled. It is work with SYMAP that has been the foundation of virtually all new work in computerized spacial analysis. The only way to achieve high resolution output is to generate very large maps out of strips (each strip being the width of the computer paper forms; and then reducing a composite map photographically to a manageable size. This can make SYMAP production runs costly. It also means that interactivity is hard to achieve, and implies one knows in advance what one is doing. SYMAP has not been the perfect exploratory tool that LCGSA had fitted it to be. One possibility which could be explored is that of using an electrostatic printer-ploter (see Appendix A-2) to replace SYMAP overstrike output with true shades. This would achieve higher resolution visually at both smaller scale and cost. Examples of SYMAP output are included in Appendix A-4.

(2) CALFORM

CALFORM was designed to overcome the resolution problem in SYMAP output and also to reduce the enormous cost associated with producing high-resolution maps. It involves a different data format and produces output by means of a computer-driven plotter. Its shading ability is limited to various cross-hatching patterns which, of course, become increasingly time-consuming as the shade darkens. It is often expedient simply to designate the value associated with a particular area with a symbol, and have it shaded or colored in by hand. CALFORM is further confined by its lack of contouring ability and its consequent dependence on the programmer to provide all points, lines, and polygons to be displayed in output. CALFORM's greatest value lies in repetitive display of information in unchanging spatial boundaries (such as Census Tracts or municipal boundaries where the data file can be input once and multiple maps made from it). With the advent of microfilm plotters (see Appendix A-2), CALFORM has become more attractive. It is now possible to produce dozens of maps (sharply-drawn and inexpensively) covering a wide range of possibilities from manipulations of the same data.

(3) SYMVU

SYMVU was designed to recapture the contouring capability lost in the development of CALFORM. It, in fact, attempted to make that contouring capability even more vivid than what was achievable in SYMAP. SYMVU abandoned shading and went to three-dimensional perspective graphics. This means that it takes an output data set and turns it into a "mountain range" where z-values of the data represent "elevation". SYMVU has two major problems: hidden lines and perspective distortion. Because hidden lines must necessarily be expunged to make the display clear, anterior high values "hide" posterior low values (mountains in front hide valleys behind). By putting the display into perspective for visual impact, SYMVU throws away any comparative value. It is difficult in a perspective display to evaluate relative levels; its only value is in making quick superficial impressions. It has little practical use in careful analysis. One nice feature of SYMVU is that its input grid is identical to SYMAP's output grid. It should therefore be easy to get a SYMVU plot if one has already done a SYMAP run. Unfortunately, SYMVU does not work on a printer, and SYMAP does not run on a plotter. To overcome data conversion difficulties, the following program was developed!

(4) POLYVERT

POLYVERT is not, strictly speaking, a mapping package. It is rather a collection of geographic data handling routines. Its primary purpose is as a manipulator of geographic base file information. It was specifically designed to make SYMAP, SYMVU, and CALFORM mutually intelligible. Perhaps the most important by-product of this creation is the clear articulation of the importance of data base design in geographic mapping and spatial analysis.

All four of the above programs have been purchased by the Division of State and Regional Planning or are available at the Department of Transportation data center. At the moment, only SYMAP is actually operational. The other three are available in listing form in our offices.

(5) ASPEX

The Harvard LCGSA is well aware that the major feature lacking from their previous product is "conversationality", or the capability for the user to interact with the system, and the program to "prompt" the user. This gap will be filled with two programs soon to be introduced. One of them ASPEX, essentially an interactive version of SYMVU; and the other is INPOM, approximately an interactive version of CALFORM. However, along with interactivity INPOM offers a data base creation facility and numerous other capabilities not available in CALFORM. It is hard to determine at present exactly what functions will be available in release number 1 of ASPEX or INPOM. It is our understanding that release 1 of ASPEX will be published sometime before the end of this year, and INPOM will follow it in about six months.

LANDSAT (Land Satellite):

This program of the Federal government combines a major new source of land use data with an information processing technique. The data are collected by means of satellite remote sensing; the data are interpreted by computer.

The program and its first satellite were formerly known as ERTS (Earth Resources Technology Satellite). Since early this year there has been a second satellite, and the program is now called LANDSAT.

The two satellites orbit the earth at approximately 130 miles. They measure light reflectance at four wave-lengths, and beam this data back to earth for processing. Different land covers have different combinations of spectral reflectance, called "spectral signatures". One satellite or the other passes over each place on the earth every nine days. The resolution of data is about one acre. Present processing technique (as described in the USGS Circular G71 developed by the U.S. Geological Survey) distinguishes three-dozen categories of land cover. Much greater differentiation is possible where agencies are willing to process tapes for themselves. The USGS is generally encouraging this to happen, and to a certain extent is ready to participate in exploratory studies.

Raw LANDSAT data can be acquired in the form of computer tapes from the NASA Goddard Laboratories, Beltsville, Maryland. If New Jersey were to acquire the capability to process tapes locally, it would have virtually inexhaustible sources of up-to-date, comprehensive land-based data.

Any such capability implemented by New Jersey would automatically benefit from Federal extensions of technique. Two major lines of development are underway. One concentrates on producing higher quality visual images from the raw data. This is the research being done at Flagstaff, Arizona. The enhanced image is then interpreted by trained photo-interpretors.

The second line of development is using computers to interpret land cover directly from multi-spectral data. Although still in their infancy, these techniques may have even more potential than the former. Some of the best work in this area is being done by James L. Wray, who came up to New Jersey from the USGS to present his recent work. An advantage of automatically-interpreted data is that it can quickly be compiled into a geographic base file, and it can readily yield specialized thematic maps.

There are two primary uses of LANDSAT data and technology for which we see a potential in New Jersey. One is the automatic production of interpreted geographic base files, just mentioned. The other is in monitoring land use changes and enforcing land use controls. Utilization in monitoring, to be wholly effective, needs both the frequency of LANDSAT coverage plus the rapidity of a local processing capability.

LUDA (Land Use Data Analysis):

LUDA is a program within the Geography Division of the U.S. Geological Survey. Its objective is to make it feasibly convenient for agencies to incorporate LANDSAT data into a general geographic information file. Interpreted information on land use derived from LANDSAT data is aggregatable to Census tracts, municipal boundaries, the outlines of major watersheds, or areas of public tenure. It will thus be possible to produce a unified base file for the entire nation. New Jersey's digitized geography should be available on computer tape early in 1976. Inasmuch as these files are in the public domain, New Jersey could acquire an initial geographical base file more than adequate for its first uses for the cost of tape reproduction (approximately $80). Estimates were made as recently as a few years ago that it would cost upward of $100,000 to create an adequate base file. It should also be noted that all boundaries in LUDA will meet national map accuracy standards, and procedures are being supplied by USGS for the subsequent modification of these files. Richard Witmer, of the USGS, has made a presentation to New Jersey on the capability of LUDA. He is the person to contact for further information on the status of LUDA activities.

MAGI (Maryland Automated Geographic Information System):

This is an operational geographic information system for the State of Maryland. It was designed by the Environmental Systems Research Institute (ESRI) of Redlands, California, and was implemented by the Maryland Department of State Planning. MAGI is probably the most sophisticated state-level geographic information system that is thus far fully operational. It consists of a versatile and updatable geographic base file for the entire state, including its water areas. It offers a multi-functional geographic analysis capability (including overlaying, contouring, and thematic map output). It is interesting to note that MAGI's output medium is SYMAP which, though around for fifteen years, is still the most flexible computerized mapping package available.

Maryland has also had two subsequent studies prepared that are of interest to New Jersey. They were done by the Earth Satellite Corporation of Washington, D.C. The first investigated the use of high-altitude aerial photography from NASA to initially classify land use. The second studied the role of ERTS remote sensor data as a source of data for maintaining and updating Maryland's land use inventory, and a way of detecting and defining areas for closer examination. Inasmuch as there are so many topographical, environmental, and cultural similarities between Maryland and New Jersey, much of their experience can be of immediate benefit.

MAGI is also a good example of what firm state commitment can accomplish in a short time. This system has been used to help make policy decisions. A thorough understanding of MAGI and the two subsequent studies should be a prerequisite to the functional specification of a coastal area information system for New Jersey.

M-DAS (Multi-Spectral Data Analysis System):

This is a complete hardware and software package for manipulating LANDSAT data. It has been developed, and is being marketed, by Bendix Aerospace Division, Detroit, Michigan.

M-DAS is a brand new product. For roughly $225,000, it provides a complete LANDSAT geographic information system. Bendix supplies an elaborate software package for editing, transforming, and cleaning LANDSAT data; and for converting it into a State Plain Coordinate grid.

Bendix Aerospace is also offering the use of its M-DAS System as a service for roughly a $1.00 a square mile. It will provide State Plain Coordinate grid cells 50 yards on the side. These data cells will be coded by USGS two-digit land cover categories, or by any other categories that can be specified from the LANDSAT multi-spectral data. The price for doing a second observation run, using the same categories but another set of tapes, would be roughly half the initial price per square mile. We have not seen M-DAS first hand, nor have we seen any of its technical documentation, but it appears to be the first in what promises to be a new line of products that combine image processing technique (RIE) with geographic base file capabilities.

The important thing about M-DAS is that it offers both a competitively priced image processing system and also a useful service function. In this latter capacity, M-DAS stands in a position intermediate between the slow Federal processing of LANDSAT done at Sioux Falls, South Dakota, and the relatively high initial start-up costs of having ones own LANDSAT processing capability. Using this kind of service at first could give an agency the valuable opportunity to get some experience in custom-making digital techniques and methodology before actually investing in an internal processing capability.

SPSS (Statistical Package for the Social Sciences):

SPSS is the most elaborate and sophisticated of the several statistical packages available for general research use in batch-entry environment. It is designed for use by non-programmers. There is very little that cannot be done with it in the way of quantitative data analysis. Its major limitation is that it is solely a batch mode package at present. Though a time-sharing environment with remote batch-entry does partly mitigate problems by simplifying job preparation and entry, batch mode still requires a relatively high degree of user familiarity.

Conversationality, or interactivity, tends both to lower the threshhold for using a system and to alter the way in which a user engages in research. The authors of SPSS candidly state that they anxiously await the pending obsolescence of their system precipitated by the development of wholly conversational programs for statistical analysis. Work is already underway to develop conversational versions of SPSS. In the meantime, the N. J. Department of Transportation has a good basic interactive package called STATPACK. Information about it is available from the DOT computing center.

The utility of DOT's STATPACK would appear to lie in exploratory data analysis in cases where discrete results can be displayed back to the user. SPSS appears to be more useful in larger statistical computations where the results will be more complex. The advantages of SPSS is that it is exceptionally well documented, thoroughly tested and debugged, and internally consistent. It offers everything from simple one-way frequency distributions to multiple uses of factor analysis and canonical correlations. To the extent that planning analysis must inevitably become more quantitative, it will call for the special tools of quantitative analysis. SPSS is capable of dealing with virtually all files, but yet is a tool accessible to users lacking extensive training in computer science. SPSS, even if not yet interactive, can be utilized effectively by anyone with a rudimentary knowledge of statistics and a few hours to spend reading the manual.

All routines in SPSS are programmed in FORTRAN. A good person to contact for further information on SPSS is Ms. Jane Li, Hill Center for the Mathematic Sciences, at Rutgers, The State University. The Educational Computer Network (ECN) of the New Jersey State college system supports and sustains SPSS.

A final word is in order on the subject of data base management systems. There is a large number of these currently on the market. Some, like IMAGE, have been developed specifically for particular computers. Others, like IBM's information management systems (IMS), are specific to very large computer systems. Still others, like **TOTAL**, come in a variety of versions appropriate for many different sizes and makes of computer. It is assumed that the major data base handling capacities of the CAIS will be supplied by the contractor who provides the central core of software. It is, however, possible that such a core would include only those routines necessary for manipulating the GBF, and that all other sorts of data base management activity would call for a resident package of file handling routines. In this case it would be wise to ascertain whether such a package exists for the particular hardware that is being considered. It is also worth noting that conversationality is as much a variable with data base handling packages as with other sorts of software. Many systems are wholly batch-mode-oriented. Others are almost entirely interactive-oriented. This is a further variable that should be weighed.

Data Base Management is a field in computer science with its own many levels of expertise. We recommend consulting with a specialist in this area before making any permanent commitment for the CAIS. Such a consultant should not be hired to design or implement a data base management system, but rather to advise and recommend a commercial package.

APPENDIX A

Section 2

Equipment Options

This section is divided into two sub-sections the first is a series of brief assertions about equipment appropriate or inappropriate to CAIS needs. The second is a more thorough definition and discussion of each of the various elements involved.

Two general sorts of questions about equipment emerge from the previous list and discussion of software system and package options. The first is a set of general questions about big computers and little computers, and how the two are best used together. The second is a set of questions about special equipment needed for a geographic and interactive information system. These two sets of questions will be discussed separately.

Computers Large and Small

The terms "computer" and "mini-computer" are no longer self-explanatory. IBM now makes physically small systems which are nonetheless full computers, and Digital Equipment Corporation (the inventor of the mini-computer) now makes systems that fill a good sized room. The meaningful threshold would seem to be mobility. If the machine can be moved by one person from room to room, then it is a mini-computer. If the machine demands a room of its own with special air conditioning or electricity, then it is a "big" computer. Put another way, can you bring it to you, or do you have to go to it. For machines that must be "gone to," DEP is well-served. The IBM 370/145 at the Department of Transportation has plenty of computing power for DEP's immediate needs. On the other hand, as a support for digitizing or the manipulation of digital images, it would be grossly inefficient to occupy an entire input/output channel of the DOT machine to transmit a few characters per minute. For this sort of use, a local dedicated mini-computer is appropriate.

The critical question then becomes that of the relationship between the two machines. Fortunately, the DOT machine is equipped to handle "intelligent" remote terminals with maximum ease and efficiency. From the other end, most small mini-computers can be equipped with the ability to interact with distant large computers as if they were terminals. It is this model of the "intelligent" remote job entry (RJE) machine operating as a conversational monitoring system (CMS) terminal that has been the guiding model for our deliberations.

Special Interactive/Geographic Equipment

Four major questions are at issue here and will be dealt with in turn.

(1) Digitizing

No matter how good the initial geographic base file is, some editing will be necessary. No file will stay useful long if updating is not possible. Some means of altering the geographic base file is thus necessary. This involves three basic capabilities: the ability to display a portion of the geographic base file at will, the ability to alter what is displayed, and the ability to compare what is altered with some map representing the "ground truth." In hardware terms, this means a digitizing table which is something on which a map gets mounted and from which coordinates can be read. Also necessary is some display device on which the outlines of the GBF can be traced. Last but not least, some means for intrepreting the latter in terms of the former is needed -- in other words, a computer. There is a handful of systems available commercially that can supply these capabilities. A minimum cost for digitizer and display combined approaches $30,000 by the time interfaces, pointers, etc., are included. The cost of the software and the computer that connect the two is harder to compute. In many cases it will be included as a portion of the system.

(2) Map Output

The product of a CAIS will be thematic maps. There are many media in which such maps can be produced; of these, the two basic media would seem to be paper and film -- the two basic techniques are printing and plotting. On paper, the print and plot options have until recently been irreconcilable. They called for two very different and equally expensive machines. With the development of the electrostatic printer/plotter, it has become possible to produce both character and line information on a single piece of paper at extremely rapid speeds. The paper moves through the printer plotter at one to three inches per second producing a diagram at a resolution of 200 dots per inch. These devices, available for $10,000 to $20,000, are only slightly more than the cost of a printer or plotter alone.

Two uses of film are worth noting. One is the direct recording of LANDSAT onto film. The output looks like a color image of the area being mapped, except that the colors are not photographic but thematic and represent interpreted land use rather than multi-spectral images. A second use of film is recording on microfilm as if it were a high-speed plotter. This work, done by Fredrick Broome at the Geography Division of the U.S. Census, has produced high-quality color thematic maps at remarkably inexpensive prices. Both of these uses of film are, however, very expensive. A microfilm device capable of producing these sorts of images costs between $80,000 and $250,000.

(3) Conversationality

The above discussion of map output assumes only a batch mode of operation. In batch mode one decides what sort of map is wanted, describes it to the computer, and the computer then puts it out as a graphic. A second sort of capability is also required and that is the ability to manipulate an image interactively in order to visualize it, point areas on it, make changes, and then see what the changes look like. As discussed in the text, there are two rather different approaches to this. One is based on points, lines, and polygons (PLP). The other is based on raster image enhancement (RIE).

PLP interactivity can be accomplished in a number of devices: The industry leader for this is Tektronix. There are a number of such devices by Tekronix ranging from small but powerful oscilloscope-like devices up through 17" television-size screens. The normal mode of interaction here is by means of a "joy stick" or X-Y wheels on the terminal itself. The motion of the joy stick or wheels is reflected in the motion of a small curser or cross mark on the screen which the operator moves to the desired location and then specifies the operation that is to be performed there. Such terminals cost somewhere between $6,000 and $12,000 depending on the interfacing and extra equipment that is required with them.

In the RIE mode, the basic capability is somewhat more expensive. The industry leader here seems to be COMTAL. The basic device consists of an industry-standard commercial color TV monitor which is rewired for digital work and connected to a device for precisely controlling the image that is produced. This device costs about $30,000 at a minimum. It is enormously sophisticated and powerful, however, and is capable of producing high-resolution images in full color with a number of different kinds of manipulation.

(4) General Interfacing

One item which has been inadequately discussed is the relationship of the special CAIS equipment to "the rest of the world" (contractors, service bureaus, other agencies, etc.). It is imperative that the system have the ability to produce all manner of output in an industry-standard medium. At the present time the only such medium is 9-track magnetic tape. It would seem obligatory that the system include a tape drive or two: one for reading and one for writing. These typically cost about $10,000 a piece.

***

This brief discussion has been intended only to pin point certain key elements within the system. It assumes a certain background of understanding about the general capabilities of large and small computers, various sorts of storage media, display devices and digitizers. A further background of these areas are supplied in the following section which is designed to introduce the reader to the basic capacities and variations that characterize and differentiate pieces of equipment.

Technical Equipment Definitions

A typical computer installation consists of many distinct pieces of equipment, known also as "hardware", which can be identified with particular operating characteristics of the computer system. A Central Processing Unit, or CPU, executes instructions and controls the transfer of information between the CPU and other units of the system. Other hardware are depositories for information (storage units and media) or serve to facilitate conversation between the user and the system (terminal, printers, plotters, graphic devices, image processing devices). The characteristics of the hardware units most likely to be incorporated into each of the CAIS alternatives are discussed in the following order:

(1) Mini-computer Processing Unit
(2) Storage Media
(3) Terminals
(4) Graphics Devices
(5) Image Processing Devices
(6) Digitizers

(1) Mini-Computer Processing Unit

The "mini-computer" of industry jargon is a stored-program digital computing machine priced under $50,000 and suitable for general-purpose applications. The "mini" (as it has come to be called) is a parallel binary processor with a 16-bit word length utilizing integrated circuits for reliable, high-speed operation. The mini's cycle time (the time required for the access and execution of an instruction) ranges from 0.8 to 1.5 micro-seconds, while input-output data transfer rates may achieve speeds as high as a million words of information per second. A single machine "add operation" (a frequent industry indicator) may take from one to three micro-seconds.

Clearly, today's minis are sophisticated computing machines capable of performing literally millions of operations (information accesses, additions, comparisons, and the like) in the time it takes a human operator to press a button. They will tolerate wide variations in their environment yet are readily programmed to perform myriad tasks. While some minis do actually resemble this "average mini" concept, the spectrum ranges from very small kinds for specific dedicated tasks to more powerful but still compact kinds that provide the flexibility and computing power equivalent to considerably larger machines of earlier generations of computers.

Considering how powerful and versatile mini-computers have become it is no surprise that they are applied in situations where automated approaches would not have been feasible in years past. The product lines of mini-computer firms are continuously upgraded or improved to satisfy user needs and the pressures of intense competition. This means that it is impossible to specify the best mini. Optimal systems configured from the current selection of minis may lose that distinction six months later as improved processors are introduced by other manufacturers. In any case, the major choice is being able to take advantage of a mini-computer in the first place, rather than marginal decisions within the mini field itself. There is much to choose from, and a system configured today should remain an attractive and viable solution for the design application in successive years.

A mini-computer of the type discussed above used as a remote extension of a larger CPU, will satisfy the information processing demands of the CAIS.

(2) Storage Media

The three widely used media of storing digitized information are punched cards (properly called "Hollerith cards"), magnetic tape, and magnetic disk. Information stored on Hollerith cards and magnetic tape is accessible in a sequential fashion only, while magnetic disk permits both sequential and random access.

The mode of accessing information in a given storage medium is limited first by the medium itself and also depends on the type of data, the size of the data files, the frequency the data files are accessed, and the frequency the data files are updated or altered. Sequential accessing of information means that the retrieval of any particular data item or record involves starting at the beginning of the file and passing sequentially, record by record, through that file until the desired record is reached. Random accessing means that the desired record may be retrieved or accessed directly, without the time loss associated with going through all preceding records. Random accessing techniques are generally more expensive to implement, yet for some applications are unquestionably superior to sequential accessing techniques.

(a.) Hollerith cards are a traditional storage medium. However, their popularity is declining, principally because new storage techniques are superior for virtually all applications. Cards are bulky, easily gotten out of order, vulnerable to heat and humidity, perpetually the victim of careless handling. Information density per card is extremely low, typically eighty characters or less. The through-put rate, the rate at which data is capable of being accessed by the computer system, is also slow, typically four hundred characters per second. The only advantage of paper cards, and it is a small one, is that they can be visually interpreted and handled manually. Hollerith cards, and their associated card reader and keypunch devices, stand to figure increasingly less in state-of-the-industry computer installations.

(b.) Magnetic tape is an attractive storage medium for most sequential processing applications. Its high-speed operation, "industry-compatible" characteristics (the ability to be shared between different computer installations), and low cost encourage wide use. Magnetic tape systems are available for all of the small computer systems examined in this survey. They ranged in price from seven to thirty thousand dollars. Record densities of 800 or 1600 BPI (bytes per inch), and recording methods that are "phase-encoded" or "non-return-to-zero" will help assure compatibility with large existing computer installations. The typical capacity of a 2400-foot reel varies between one and five million characters of information (where a character is equivalent to a byte), while information can be accessed at speeds up to 30 thousand characters per second.

(c.) The magnetic disk is an attractive storage medium for many sequential, and virtually all random access, data processing applications. It is a very high speed device, featuring information access rates of up to a million characters per second. The individual disk pack will accept from 256-thousand to over 80-million characters of information. Magnetic disks are commonly bundled into a system of disk packs providing billions of characters of information storage. Disk packs vary in price from $10,000 to $60,000. While the magnetic disk is faster than magnetic tape and provides for random access of stored information, it is more expensive and is not industry-compatible.

(3) Terminals

A terminal is a device to facilitate conversation between the user and the computer system. The terminal is the way that the user controls features of the computer system available to him and instructs, or "programs", the system to operate on data. The system in turn can respond back to the user's commands and instructions via the terminal.

Not all terminals provide a printout of what transpires between the user and the system. Those that do, known as "hard-copy" terminals, print information at rates varying from 10 to 30 characters per second and utilize a standard keyboard. Prices range between $2,000 and $5,000. The "soft-copy" terminals, those which provide no physical printout of the terminal session, display information on a television-like CRT (cathode ray tube) at rates up to several thousand characters per second, commonly in black and white (although color displays are available). The soft-copy terminals will also accept information from a standard keyboard, and cost between $2,000 and $9,000.

Each terminal type has distinct advantages and disadvantages. Soft-copy terminals are quiet, require little routine maintenance (like paper changing), are very fast, and are popular in office locations. Hard-copy terminals are noisy and slower, but do provide a permanent copy of the terminal session for annotation, presentation, or discussion. Different applications dictate having one or the other or even both. Although it is possible to constrain applications to just one terminal, it is probably desirable to have a high-quality version of one and an inexpensive version of the other considering the low contribution their prices make to the cost of the total system.

(4) Graphics Devices

Graphic displays provide pictorial representations of spatially-organized or graphically-intelligible data. In the sense that they provide communication between the user and the computer system they may be considered specialized kinds of terminals, and as such are available as either "hard-copy" or "soft-copy" devices. Included in the category of soft-copy devices are raster scan displays, storage CRTs, and refresh graphic displays. Hard-copy devices include the traditional printers, the drum or table plotters, the more recent electrostatic printer/plotters, and computer output microfilm (COM).

(a.) Raster scan displays use a technique equivalent to the standard television picture tube for its display. The image is formed by the energy of a scanning electron beam striking the phosphorescent coating of the tube's inner surface. Since the electron beam scans the surface of the tube a line at a time, raster scanning devices tend to be considerably slower than other display devices.

The raster scan device is not suitable for the real-time generation of flowing line graphics or user-modified displays. The device finds application in displaying unchanging stored images and in some image processing applications such as primary spectral interpretation of ERTS digital tapes.

(b.) Storage tube devices also utilize a cathode ray tube (CRT) but operate in a way significantly different from raster scan techniques. Information is sent to the device which indicates the location of a point within the picture being displayed. An electron beam then strikes precisely that point on the CRT's surface, causing the surface at that point to floresce. The brightness endures for several seconds, during which the device excites other such points on the CRT's surface to complete the image.

High resolution, effective graphics manipulation, and high-speed operation characterize the storage tube device, which ranges in price from $5,000 to $9,500.

(c.) Refresh graphic displays operate much as storage tube devices, but each display point persists for a significantly shorter time than is typical of a storage tube device. This means that refresh graphic displays offer very high speed and high resolution, in addition to permitting real time interaction with the displayed image. By using a light pen, the user may actually draw pictures on the screen and have them maintained or altered by the computer system. This function is certainly advantageous for some applications but must be weighted against the demand placed upon the computer system. Since each point persists for only a very short moment, the computer system must maintain, or refresh, every point very frequently, which means that the overall performance of the system is degraded. In addressing this problem, companies are developing refresh graphics displays with their own processor, thereby making the device less demanding of the main computer system. These devices have yet to be investigated in detail.

Depending on the complexity of the device, refresh graphics displays will cost from $3,000 to $25,000.

(d.) Of all the hard-copy graphics devices, the impact printer is familiar to everyone. For many years it provided the only means of computer graphic output. The standard character set of the printer, and its fixed implicit grid, was used to approximate shapes and lines. Gray shades could be achieved to a certain extent by overprinting characters. The deficiencies of this approach to graphics are manifest. Were it not that some of the most useful computer graphics programs that have ever been produced (such as SYMAP) happen to have been developed during this era, and consequently require printer graphics, the utilization of standard printers as graphic output devices would have long since become an extinct practice.

They furthermore fail to achieve the degree of visual clarity necessary for effective graphic communication. The implicit printer grid lacks the programmable resolution needed for clear "up-close" graphs or displays. As a consequence, impact printer graphics are often produced at large size, spliced together, and displayed at a distance sufficient to make the pattern emerge out of the constituent individual characters. The problem is akin to not being able to see the pattern of the "forest for the trees". The problem is compounded in that printer characters do not fill their grid positions but leave a network of prominent white framing lines over the display. Unfortunately, the pigment saturation of impact printers is so light that they cannot be displayed at a distance without some enhancement (such as going over character patterns with felt-tip markers in different shades or colors). The net result is that impact printer graphics are not satisfactory up close or at a distance!

(e.) There are many alternate ways to produce hard-copy displays without hammering ink on paper. These techniques use chemicals or electronics, and include ink jet, optical, magnetic, thermal, xerographic, electographic, and electrostatic methods. We have been partial to the electrostatic technique because it can take advantage of extant impact graphic programs, yet has numerous advantages over impact graphics. First of all, there are no moving parts in the image-making process except for the paper transport system. A dielectric-coated paper is passed over the writing head where it is charged with a pattern of electrostatic "dots" as dictated by the data source. Liquid toner is then applied to the paper to develop the latent image which, when dry, creates a permanent, archival-quality hardcopy.

Other major advantages of an electrostatic image-making technique are high output speed, low initial cost, low operating cost, high-contrast displays, high-quality plotting capability, and the ability to accommodate a wide variety of letter fonts and sizes. Leading electrostatic printer/plotters offer resolutions of 100 to 200 dots per linear inch and can print from 200 to 1200 lines per minute depending mostly on the CPU. Clarity, speed, and resolution of this degree offer a major breakthrough for computer-generated displays and computer-produced documentation.

A printer/plotter of moderate capability sells for about $10,000. This is roughly $6000 less than comparable impact printers. If a printer/plotter avoids the need for both an impact printer and a pen plotter, savings of $10,000 to $15,000 might result.

(f.) Plotters are the next most traditional approach to computerized graphics after impact printers. Though great technical innovations have been made in plotter equipment, they still essentially consist of a writing stylus moving in serial manner over a writing surface.

The writing stylus is frequently a pen or inking mechanism which produces a crisp, precise line. Felt-tip pens are also sometimes used to achieve a wider line with higher display visibility.

The graphic surface can be paper, acetate, mylar, or specially-prepared cartographic material. In nearly all cases, the writing medium is fed from a continuous roll. This eliminates frequent paper-changing and also enables variability in the length of the graphic. The width, of course, is fixed by the size of the paper roll.

The writing surface is mounted on either a revolving drum or a flat bed. In the case of drum plotters, the stylus can only move in the lateral direction. Line-drawing in the perpendicular direction is accomplished by the drum itself which advances and reverses the writing surface under the laterally-moving stylus.

In the case of flat-bed plotters, there are two independent advancing mechanisms -- one at right angle to the other. The simultaneous operation of the two mechanisms causes the stylus to draw the vector sum, or resultant, line. If the actual configuration of equipment is synchronous, a free-flowing line can be achieved. It is more common, however, for the mechanisms to alternate instantaneously. This causes a "stepped" line to result. Depending on the smallness of the increment, the stepped nature of curved lines may approach imperceptibility.

The finer the resolution, the more expensive the plotter, however. This is true of both drum and flat-bed plotters. Often, a machine with relatively coarse incrementing is gotten in order to keep the cost of this equipment down to a reasonable fraction of total system cost. This may mean that a plotter that is wholly satisfactory in regular analytic work, may prove too crude when the "finished" quality of presentation graphics are needed. The right balance between cost, speed, and plotter resolution is difficult to triangulate. A perfect solution may not be possible. The versatility achievable with one machine that characterizes the electrostatic plotter, is not typically a property of drawing plotters.

(g.) Computer-output microfilm (COM) is a competitively-priced alternative to plotter graphics. Instead of drawing actual lines onto physical paper, COM records photographic lines onto microfilm. A major advantage of COM is its extremely low unit costs.

It is, for example, feasible to do many inexpensive microfilm "plots" of time series data (either historical or forecasted), and then project them in sequence. This achieves a computer-generated "animation". Instead of animating a story-line about persons and characters, a movie can be made of geographic change.

Perhaps, an even more successful approach would be to include a COM sequence as a segment of a traditional narrative movie.

Computer animation is a rapidly-developing front in graphic analysis. The same line of argument that leads us to recommend interactive graphics and visual display analysis, has also led us to recognize the tremendous communications potential of COM. A major role in planning and public information methodology could be taken by COM.

The one major drawback presently is high initial equipment cost (anywhere from $100,000 to $400,000). Although it is true that this is recouped gradually in low operating costs, this applies most convincingly to established graphics systems where the volume of use and output can accurately be predicted. In our opinion, serious consideration of COM can safely wait until equipment costs go down, techniques are refined to be more user-oriented, and the exact nature of potential use is fully discussed. It should also be noted that there are established firms offering good services very near Trenton.

(5) Image Processing Devices

Image processing devices are relatively new additions to the computer peripheral arket. This type of equipment finds applications in the processing of information such as the satellite imagery available from the Earth Resources Technology Satellite (ERTS) and the subsequent LANDSAT programs. Remote imagery data is received as selected points of information, or picture elements ("pixels"), which can be reassembled in proper spatial orientation to form an image. The subsequent interpretation of the image was explained in the body of the report.

Image processing devices not only serve in the creation of an image from a selected source, but also facilitate the display, enhancement, or alteration of an image, in gray shades or in color. The devices are extremely sophisticated, consisting of a dedicated processor, extensive special software, and also specially-designed peripheral equipment.

The cost of the equipment, which operates "stand-alone" (independent of external systems), ranges from $32,000 to $80,000, Because of the high price tag, it is sensible for low-volume users to go to an outside service for image processing, rather than purchase or lease expensive equipment.

(6) Digitizers

The limiting factor in the preparation of a geographic base is the acquisition of geographic location data in the proper digital form. Somehow, forms contained on maps must be converted to a format in which computers can use them. This process of converting traditional mapped information into a computer-storable and computer-readable format is called "digitizing".

The traditional technique for digitizing a map consists of translating lines (or curves considered to be a series of short line segments) into ordered sets of Cartesian (x,y) coordinate pairs. Each coordinate pair represents a mapped point in a machine-readable and machine-storable form. A chain of such points, given to the computer in order, constitute instructions on where to plot or print a line or an area. Connected, these instructions and groups of coordinate pairs can reconstitute all the locational information of the original map.

Manual digitizing, converting mapped information by hand into geocoded format, is a very slow process. Because it involves continuous hand work and great concentration, it is also subject to decreasing returns to scale, and to extensive systematic and random human error. The Division of State and Regional Planning has had some firsthand experience with manual digitizing methods gained during a prototype development of a land tenure geographic base file. Based on observed unit rates, it was estimated that three to four person-years would be required to prepare the basic geographic file for the State using manual methods. It was also estimated that the rate at which data base elements could be corrected by manual methods was probably slower than the rate at which changes occur. For any geographic base file as ephemeral as land tenure, land use, or geographic land cover, it is obvious that automated technique is needed. Only in situations where the geographic base can be expected to remain stable for many years would a wholly manual digitization approach be feasible.

The two major types of digitizing equipment are electro-optical devices and the more traditional electro-mechanical devices. The electro-optical devices utilize optical techniques to interpret point and line information. The electro-mechanical devices use mechanical analogs such as tracking arms to measure point and line information.

(a.) Optical Digitizing Devices

The optical devices used for digitizing are capable of very high speed operation, but are extremely costly. One of the most sophisticated of such devices, the I/Ometrics machine in use by the U. S. Bureau of the Census, cost about $750,000. Needless to say, for devices of this cost and speed magnitude, work is done on a contract basis.

From what we have heard, optical digitizers are capable of producing many hundred thousand points of digitized data per day, including all corrections and rectifications, for under a hundred dollars. A number of leading manufacturers and service bureaus have recognized that, in many cases, although the need for massive digitizing may initially be strong, there is no long-term need for the high capital expenditures and related personnel overhead of the most sophisticated digitizing processes. A growing number of firms now offer a full range of digitizing services. This is an extremely attractive way to close the gap between the information soon to be available in the public domain from the LUDA program of the U. S. Geological Survey, and the more detailed information base requirements that an operational State planning and permit review system can reasonably be expected to produce. A contract approach can save much valuable time at the beginning of an information system for tasks more important than the routine infilling of the geographic data base. It might also be convenient at a subsequent time to extend the geographic base from merely coastal counties, for example, to coverage of the entire State. Contract digitizing to a firm with the specialized software and personnel needed for geo-base expansion is a rapid and efficient way to accomplish this.

(b.) Mechanical Digitizing Devices

For ongoing correction, editing, and maintenance of a geographic base file, an optical digitizing device (or the services of a speialist firm that has one) are not necessary. In regular week-to-week operations, an electro-mechanical digitizer will be perfectly satisfactory. There are two essential types of mechanical digitizers: the fixed-arm digitizer and the floating-cursor digitizer.

The fixed-arm device is a relatively inexpensive machine costing a few thousand dollars. The device is also characterized by relatively slow operation--usually only a couple hundred points per day, depending on the ability of the operator. Electro-mechanical devices can, in principal, be either on-line or off-line, though few CPU users would appreciate the system degradation that slow-operating fixed-arm digitizers would cause.

The floating-cursor digitizer is similar in result to the fixed-arm device except in the important instance of faster data acquisition speeds. The floating-cursor device gives the operator far more freedom and tends to increase personnel productivity as well. To achieve this increased ease of operation and faster data acquisition, one finds that typical devices cost around $16,000.

The actual choice of a specific digitizing device depends on a detailed appreciation of the probable nature of use to be made of the resulting information system. The more geographic detail desired, the greater the extent of interactivity wished, or the faster the inclusion of current data--the stronger the case for having a free-cursor kind of electro-mechanical digitizer instead of the fixed-arm variety. In any case, it is recommended that only the needs to maintain, alter, and edit the geo-base be met with an in-house digitizer. The initial creation of the CAIS data base should take advantage of the economies of scale of optical digitizing devices as much as is feasible.

*

APPENDIX A

Section 3

Conclusions

The most striking discovery of our investigation into the State of the Industry in geographic information system design was that the plausible options are strikingly limited. More precisely there are thresholds. This is to say that a wide range of equipment options exists but that in terms of actual system capabilities there would seem to be only three or four action options open to DEP.

One option is of course to do nothing. One could hire a million "Kelly Girls" to handle the various aspects of CAFRA (monitoring, planning, public participation, enforcement, etc.). This option can hardly be taken seriously.

There are however three options of system configuration that could be considered. We feel strongly that only one of these is a viable option. But it is necessary to identify the others if only to dismiss them.

It would be possible for a relatively small outlay in equipment and staff (perhaps $35,000 plus three people) to create a prototype geographic information system for the coastal zone. It would have to be housed entirely at the Department of Transportation, and it would consist solely of commercially-available software and administratively-available data. The system could depend entirely on DOT's 370/145 for computer power, and operate permanently on a static geographic base file. This system would be little more than an illustration of what might eventually be accomplished. Each individual step in its development would have to be devised, tested, and implemented by its three staff members. In our opinion this option would be essentially an expensive toy with no chance of being able to do enough soon enough to actually aid in the implementation of CAFRA.

A second option more difficult to dismiss is that of acquiring the necessary equipment and staff, working entirely with commercial packages. It would make it possible to have the staff at DEP, but would give them no in-house computing power. It would also require that all special purpose software be designed from scratch. This option, costing around $90,000, would include a digitizing system as well as an RJE terminal with printer/plotter output and the software necessary to implement an updatable data base such as that of LUDA. This option basically yields the mere ability to produce thematic maps on a known file with known data. This is far short of the versatile, interactive, and exploratory analysis capability that is needed. The ability to form "what if" questions calls for a large step -- one we feel is also a defining step of the system.

The third and most expensive option is that which we have proposed. We have estimated that to acquire all equipment and software necessary for a fully-interactive geographic information system would cost between $200,000 and $250,000. Such a system would be based on a dedicated mini-computer at DEP which would handle all editing and interacting functions but would rely for all major data processing on the large computer at the Department of Transportation. This does not quite include all possibilities since it assumes that <u>RIE</u> capacities could be deferred until the industry prices come down below the current levels. Nevertheless, the state of the industry is such that a completely functional (and now conventional) <u>PLP</u> geographic information system can be had for a price that fits within DEP's budget. We feel that anything less than this ignores a major opportunity.

APPENDIX A

Section 4

Examples of Computer Graphics

Examples of conventional printer graphics, plotter (mechanically-drawn) graphics, electrostatic printer/plotter graphics, computer-output microfilm graphics, and image processing graphics are available at the Department of Environmental Protection or the Division of State and Regional Planning for examination. In most cases, the graphics do not reproduce well so there is no point in including them in the copies of this study being distributed for review and comment.

C A L E N D A R

DSRP has no experience as a systems contractor -- and thus no basis for assigning quantitative time-frames to the above schedule of implementation activities. But the sequencing of those activities has an internal logic, from which certain conclusions can be drawn.

- IF the conclusions of this report are accepted with little or no modification by both DEP and the Bureau of Data Processing Management (Treasury); and

- IF the staff nucleus is thereupon assembled by assignment of current employees; and

- IF consensus among all affected parties is quickly reached as to the sequence and priority of CAIS performance objectives;

THEN: it is not unreasonable to expect the publication of specifications for bidding by September, 1975.

- IF the response to the RFP by the vendor community meets all specifications -- including those of system integration; and

- IF that response includes no great surprises, thus minimizing the time needed to study bids and configure a system;

THEN: a contract award sometime in December 1975 may be anticipated, with equipment delivery some 60 days thereafter.

- IF, in the meanwhile, a separate staff commitment has been made to allow the design of data bases and collection of data to begin as soon as the RFP is published; and

- IF a third CAIS task force has been similarly freed to design and develop the first system application; and

- IF six months (Oct - March) proves to be enough time to complete these interlocking tasks; and

- IF the purely mechanical aspects of the system can be shaken down within 30 days of delivery (i.e., during March);

THEN: each of the three main subsystems of CAIS will be ready for integration by April, 1976. The support system will have been installed and tested; the two data bases will be ready for implementation; and the premier system application will have been designed and prepared. This leaves the final quarter of FY76 for the "maiden run" of a completed CAIS, which is consistent with the calendars proposed in CAFRA and in the Second Year Program under CZMA.

1: REPRESENTATIVE GEOGRAPHIC INFORMATION SYSTEMS

System Name and (system acronym)	Place	Mandate	Administration	Contractor	Date of Implementation
Land Use Natural Resources Inventory (LUNR)	New York State	Governor	Planning Division (department of state government)	Cornell Univ. and Harvard Lab. for Computer Graphics and Spatial Analysis	1969
Arizona Trade-Off Model (ATOM)	Arizona	Governor	Department of Economic Planning (department of state government)	Battelle Institute-Columbus	1972
Louisiana Environmental Management System (LEMS)	Louisiana	Legislature	Joint Committee on Environmental Quality	Geoscience, Inc.	1973
Broward Impact Zoning System (BIZS)	Broward County, Florida	County Government	Local Planning Board	IBM (Federal Systems Division)	1974
Maryland Automated Geographic Information System (MAGI)	Maryland	State law	Department of Planning (department of state government)	Environmental Systems Research Institute	1974
Calista Automated Land Management System (CALMS)	Alaska Central Area	Federal law settling native land claims	Calista Alaskan Natives Corporation	Geoscience, Inc. (now part of Calista Corp.)	1975

Appendix C:

2: SYSTEM PURPOSE AND COSTS

System	principal purpose	principal client	secondary purpose or clients	development cost	annual cost of operation	staff
LUNR	archival land use inventory only	state planners	regional and local planning	$1 million	$50,000 per yr.	5
ATOM	economic vs. environmental trade-offs	state planners	help developers locate sites	$250,000 contract + $25,000 in-house	system never made operational	---
LEMS	assessing impacts of types of development	regional planners	none yet	$250,000 contract + $200,000 in-house	$100,000 per yr.	6-8
BIZS	impact model to assess policy	local planners	too new to say	$650,000	too new to say	3
MAGI	planning analysis and research	state planners	system too new to say	$200,000 contract (+ large in-house commitment)	too new to say	6-8
CALMS	land selection by multiple criteria	Alaskan natives corp.	natural resource identification	$350,000	too new to say	3

Appendix C: 3: SYSTEM CHARACTERISTICS

System	data base	physical area covered	resolution of data base	size of data base	type of central processing unit (CPU)	CPU access moc
LUNR	1-km grid	100,000 sq.mi.	half grid cell	14 million bytes	IBM 370/158	batch from remote terminal
ATOM	6 mi x 10 mi grid	more than 100,000 sq.mi.	half grid cell	20 million bytes	CDC 6400	batch from remote terminal
LEMS	grid and polygon data base	not yet complete, so it varies	varies	incomplete	CDC 6600	batch from remote terminal
BIZS	DIME file	650 sq.mi.	city block	incomplete	IBM 370/135 (batch)	IBM 360/40 dedicated
MAGI	0.6-km grid	12,000 sq.mi.	half grid cell	5 million bytes	IBM 360/50	batch from remote terminal
CALMS	unknown (proprietary)	30,000 sq.mi.	unlimited	20 million bytes	Nova (mini-computer)	dedicated machine

Appendix C: 4: SYSTEM CAPABILITIES

stem	modeling	overlaying data	factor weighting	contouring routine	graphic display	inter-activity
LUNR	no	yes	no	yes	printer graphics (SYMAP)	no
ATOM	no	yes	arithmetic	no	printer displays	iteration
LEMS	yes (specific models)	yes	no	no	plotter graphics	no
BIZS	yes	CRT (cathode ray tube)	yes	no	printer & plotter graphics	yes
MAGI	yes (general capability)	yes	arithmetic	yes	printer graphics (GRIDS)	iteration
CALMS	yes (general capability)	yes	arithmetic & logical	no	CRT, plotter displays	conversational

5: DEVELOPMENT CHARACTERISTICS

System	Staging	Plans for future	Comments
LUNR	one-shot implementation of full system	to be replaced by new archive with more analytic capability	vanguard geographic information system; huge archival effort still useful to planners at all levels of govern ment; replacement by "LRIS" in FY76.
ATOM	prototype only using dummy data	plans to cut the capability back to just MCD data, modelling routines, for economic & demographic analysis only	first to focus on policy examination and to emphasize models; has subsequently abandoned grid-cell resoluti and all environmental data; bad contract let consultant off before finishing system.
LEMS	prototype using data for test area and nine case studies	smaller CPU; optical-scan digitizing; more general modelling; new data types	earlier work of developers of CALMS; noteworthy for variet of data and firm focus of capabilities on assessment needs of decision-makers; al software in public domain.
BIZS	one-shot implementation of full system	unknown	system just installed; very advanced analytic programs and high resolution in small area; unusually expensive on a unit-area basis.

(continued)

5: DEVELOPMENT CHARACTERISTICS (Cont'd.)

stem	Staging	Plans for future	Comments
MAGI	one-shot implementation of full system	use earth satellite imagery; new models and applications	most advanced system in public sector; represents long-term effort to bring the state of State planning to the state-of-the-art level; begun using earth satellite land use source data.
CALMS	full implementation of partial system	extend area; new data types; contouring of earth satellite data	designed to be at the state-of-the-industry; minicomputer base will mean great saving in cost of operating; most functions contained in one software package (including digitizing capability).

Memorandum of Understanding

It is agreed that the Division of State and Regional Planning, beginning December 9, 1974, will perform the Scope of Services, attached and made a part of this Memorandum of Understanding, for that portion of the program called "Revised FY 75 Work Program for Development of New Jersey's Coastal Zone Management Plan", and for performance of this Scope of Services the Department of Environmental Protection will commit $25,000.

On a monthly basis, after submission of documented costs, the Department of Environmental Protection will reimburse the Division of State and Regional Planning for those costs through a properly prepared Certificate of Debit and Credit. Any services under this agreement may be sub-contracted only as agreed to by both parties. All travel costs charged against this project will be consistent with the New Jersey Department of Treasury Travel Regulations.

Both parties reserve the right to extend this Memorandum of Understanding subject to a revised Scope of Services to be prepared on or about January 10, 1975.

If for any reason it is decided by mutual agreement to discontinue this relationship, any costs incurred to date will be reimbursed to the Division and all documents prepared will be made available to the Department of Environmental Protection.

December 6, 1974
date

Edward B. Feinberg

Rocco V. Ricci

for the Department of Environmental Protection

December 6, 1974
date

Richard A. Ginman

Richard A. Ginman, Director
Division of State and Regional Planning
for the Department of Community Affairs

Appendix D:

6 DECEMBER 1974

WORK PROGRAM

Introduction

The development of a Coastal Area Information System (CAIS) involves two major interrelated activities:

- the organization of relevant data into a system of automatically processable geocoded files;
- the design of a computer-assisted planning system based on automated geographic analysis.

Fundamental to both these tasks is the specification in operational detail of just what it is that the completed CAIS will be expected to do, and for whom, and with what ease and reliability.

The development of such precise system performance specifications will require the evaluation of a large number of separate subsystem operations in terms of four basic parameters:

- the nature of the data universe involved (small enough to keep on line? drawn from many different files?)
- the projected expertise of the operator (computer professional? concerned citizen?)
- the location and sophistication of the equipment involved (in-house or remote? dedicated or shared?)
- the desired response time between inquiries (conversational or batch mode?)

The answer to these questions will almost certainly be quite different for different parts of the system. Preliminary work of this kind suggests that inconsistencies are all but unavoidable. It will be necessary to compile a series of systems alternatives, and choose among several groups of potential system performance goals.

It being axiomatic that all the desired performance goals can never be achieved within the constraints imposed by limited resources, the most critical aspect of the design process will be the assignment of priorities across the catalogue of program goals, and the negotiation of trade-offs among these competing claims on project personnel, funds and equipment. It is as a consultant in these areas that Division of State & Regional Planning proposes to make its primary contribution.

The primary goal of the FY 75 work program is to resolve such major design decisions, and produce a detailed set of functional specifications that are consistent with (1) the CAFRA mandate, (2) the federal Coastal Zone Management Act, and (3) the available and projected resources for CAIS system development. This system design will take the form of a detailed work program for FY 76, leading to the implementation of an operational Coastal Area Information System.

In addition to an operational description of system performance goals, this work program for system implementation will demand the following prerequisities:

I. organization of relevant data for system usability

II. evaluation of available software and specification of needed additional capabilities

III. determination of best hardware configuration

IV. development of a (crude) system prototype as a basis for evaluating alternatives and estimating implementation schedule

These four items form the outline which the Division of State and Regional Planning proposes for the remainder of FY 75.

Tasks

I. DATA

A. Assist the OEA in the development of an inventory of data sets being studied under Sections 2.1 through 2.4 of the Revised FY 75 Work Program for Development of New Jersey's Coastal Zone Management Plan (RWP), dated November 1974, for inclusion in the Coastal Area Information System (CAIS).

The product of this phase will be a documented evaluation of data sets located by OEA under the RWP. Elements considered in the evaluation will include:

- data storage medium
- record layout
- ease of input standardization
- geocodability
- cost of acquisition/implementation
- implicit machine requirements
- evidence of completeness, accuracy and updatability
- relationships among data sets, relevant to data base management concerns

B. Assist the OEA in the design and implementation of a format "for acquisition of environmental inventory data from a variety of sources" appropriate for use with a preliminary CAIS, as defined in section 2.6 of the RWP.

II. SOFTWARE

Identify and evaluate the available software packages applicable to CAIS, and specify those software capabilities that will have to be developed especially for CAIS.

The ultimate product in this phase will be a complete set of system performance specifications for an implemented CAIS. For the reasons set out in the introduction, it is not possible to specify at this time the nature of this product with any precision. Many decisions remain to be made as matters of DEP policy in this area. These decisions call for the articulation of alternative performance goals, and the implications of different actions at key decision points.

The immediate product under this work element will be a documentation of the major options available (whether through government sources or on the private market) in each of two areas:

- storage/retrieval systems (DIME, LARS, IMS, etc.)
- mapping/analysis systems (ERTS, LCGSA, APL/CMS, etc.)

The activities under this task and the next will be simplified and accelerated by participation in the coming International Conference on Automation in Cartography. This forum, hosted jointly by the American Congress on Surveying and Mapping and the United States Geological Survey, will be held at the USGS National Center in suburban Washington, D.C., December 9-12. This Conference presents fortuitous opportunities both to evaluate alternative systems and to examine potential hardware configurations. Recognizing the time constraints under which CAIS development must take place, attendance at this Conference is considered vital to the work program.

III. HARDWARE

Identify the major decisions to be made regarding CAIS equipment and systems facilities; develop criteria for (1) selecting the most suitable large-scale ADP installation for CAIS implementation, and (2) deciding when to acquire in-house equipment, and when to purchase the equivalent services.

The product of this work element will be a detailed analysis of the costs and benefits to DEP of implementing the CAIS in each of several configurations. This analysis is conceived to be an ongoing activity, rather than a single document to be submitted at the end of the work period (although a history of the analytic effort will be provided at that time). It is anticipated that many steps in the analysis will be highly context-dependent, i.e. many hardware items will be attractive only as elements in a particular configuration. The primary task of this analytic effort is the articulation of these alternative configurations and their several implications, in order to assure the wisest possible investment of resources at the time of actual system implementation.

IV. PROTOTYPE

Produce and demonstrate a prototype geographic analysis system, for use as a base-line in comparing alternative approaches, projecting the CAIS implementation schedule, and evaluating potential techniques of system application.

This prototype is intended to test the contention that all CAIS files can usefully be related to the New Jersey State Plane Coordinate System, and is expected to include the following elements:

- geographic base files of 1970 US Census Tracts;
- data gathered as part of the USGS CARETS project;
- data gathered under the mandate of the Wetlands Act;
- procedures for producing automatic graphic analysis of any data that can be identified with one of the above data sets;
- procedures for creating and editing geocoded files;
- a method for identifying any arbitrarily defined "area of critical concern" in terms of these data;
- demonstration of all the above elements in systemic operation, including explicitly drawn analogies to prospective applications of CAIS.

Schedule

The four areas detailed above represent a general overview of the proposed DSRP effort to be funded under this memorandum. An explicit element in each of these areas would be the preparation of detailed work plans covering the activities to be conducted during a given calendar period. These work plans will be expected to include:

- definition of measurable objectives;
- assessment of resource needs (manpower and funds);
- identification of those program elements which can proceed independently of each other--and of those which rely on output from some other element as necessary input;
- description of a series of fallback positions (if we aren't at point X by time Y, then what?);
- a plan for detecting problems in advance (i.e. how to avoid having to take one of the fallback paths);
- assignment of functional and administrative priorities to various work elements, as a basis for making critical resource decisions in mid-stream.

For the reasons discussed in the opening section of this memorandum, it is not realistic to attempt to project such precise work plans very far into the future. Rather, it seems most prudent to chart the activities in each work area from one major decision point to the next, as each of the basic decisions is confronted and resolved.

We therefore propose that work begin immediately and simultaneously on tasks I through IV, with a requirement that DSRP produce a detailed work plan covering the period through February no later than January 10, 1975. A similar plan for the period covering March and April, 1975 will be due on February 28. The final two months of the fiscal year will be covered in a plan to be submitted April 28, 1975.

COASTAL AREA INFORMATION SYSTEM: A Proposal

(Division of State and Regional Planning
18 March 1975)

The Division of State and Regional Planning has been operating as a technical assistance arm of DEP's Office of Environmental Analysis since December 9, 1974. We have analyzed and evaluatedtechnical issues relating to implementation of New Jersey's Coastal Zone Management Plan.

Our investigation focused on what kind of system could be put together "off the shelf" using the "state of the industry" rather than the state of the art. This assumption appears justified. We found that the present state of technology exceeds the ability to use it.

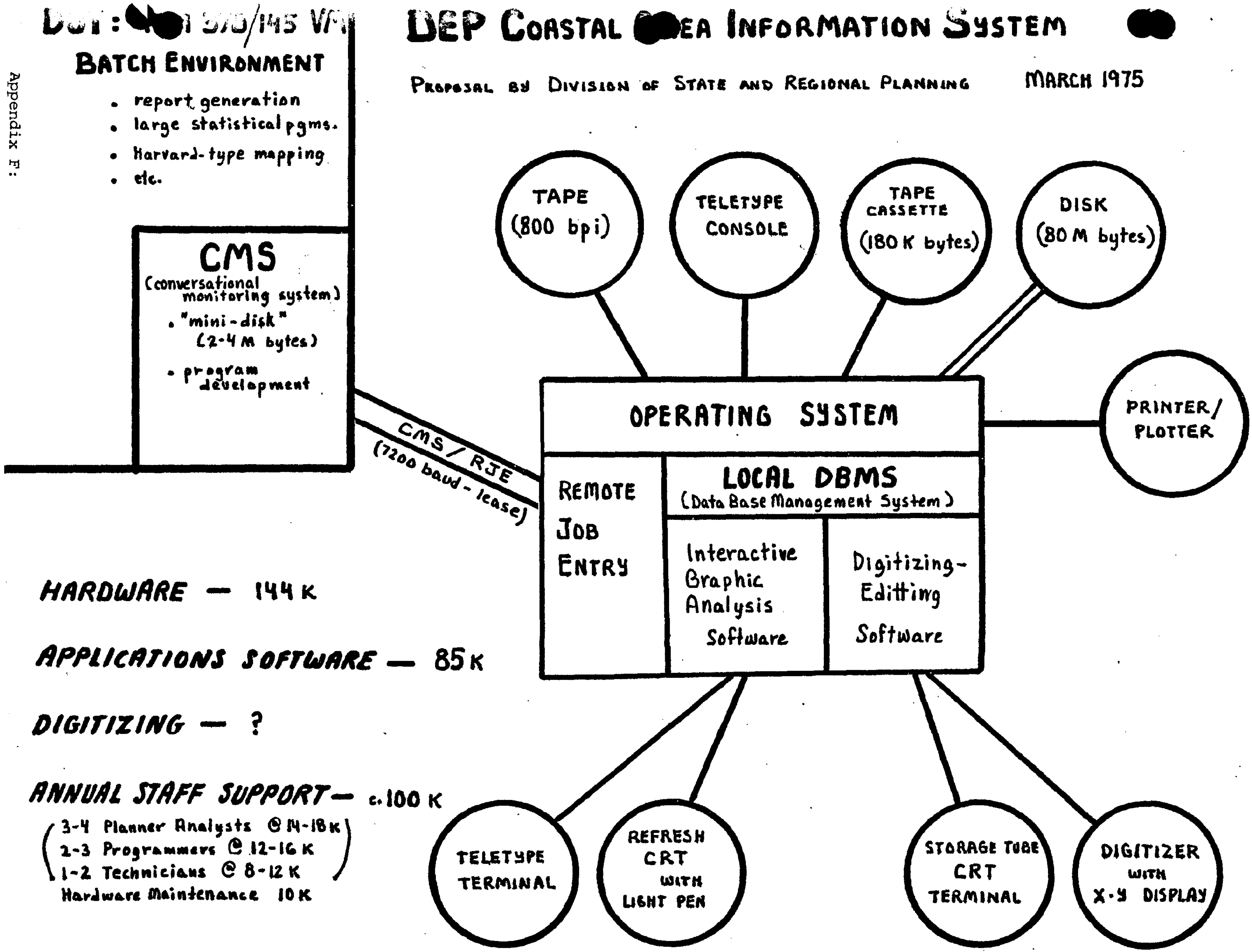
DEP has a problem to solve: it is not how to design the best of all possible geographic information systems, but how to manage the Coastal Zone within the mandate of the legislated timetable. If purchase orders go out in June, a system can be brought up by October. This would bring the powers of automation and computerization to the routine manual tasks of large file maintenance. The chief features of such a system would be graphic display and user interaction. This substantially speeds up the processes of evaluation and analysis. A Coastal Area Information System would:

- maintain an inventory of environmental resources and other data for the monitoring and management of economic development in the CZ;
- evaluate alternative sites or routes for proposed development in terms of variable criteria;
- project secondary impacts of proposed policies or developments; and
- serve as a basic analytic tool for the development of a long range Coastal Zone Management Plan.

We have concluded that a more than adequate system can be configured from hardware, applications software, and extant geographic base files in under sixty days. We estimate the cost of hardware at $144,000; possibly nothing for digitizing; $85,000 for applications software; and in the neighborhood of $100,000 for annual staff support (much of which is already on board and budgeted).

DEP COASTAL AREA INFORMATION SYSTEM

PROPOSAL BY DIVISION OF STATE AND REGIONAL PLANNING — MARCH 1975



***HARDWARE* — 144 K**

***APPLICATIONS SOFTWARE* — 85 K**

***DIGITIZING* — ?**

***ANNUAL STAFF SUPPORT* — c. 100 K**

(3-4 Planner Analysts @ 14-18 K
2-3 Programmers @ 12-16 K
1-2 Technicians @ 8-12 K)
Hardware Maintenance 10 K

Appendix F:

DETAILS ON CONFIGURATION

I. Geographic Base File (GBF)

A. As a part of its LUDA project, the USGS is preparing a national GBF that will depict, to national map accuracy standards:

- . land use (per USGS circular #671, level II)
- . watersheds (perimeters only)
- . 1970 Census tracts and political boundaries
- . land tenure (Federal, State, private)

Stable-base transparencies of these maps, made to overlay on the USGS Topographic series (1:250,000), will be available for New Jersey by this Fall; merged and edited digital files are promised by the end of the calendar year.

COST: $100.

B. There is now a variety of service agencies specializing in the preparation of digital GBFs. In particular, i/o Metrics, the firm preparing the LUDA files, is available for contracted extensions of that work. Prices in this market vary dramatically, as do performance ability (no bidder was able to meet the RFP standards of the LUDA project using manual techniques).

C. To update the GBF and provide for the inclusion of problem-specific details, some in-house digitizing capacity is required. Hardware options range from $1200 table top mechanical models to $150K programmable-laser scanning gear (the basis of the i/o Metrics system). A useful middle ground is occupied by free-cursor electronic digitizing tables supported with interactive file-editing software written for a minicomputer. Such systems combine great accuracy with enough speed and ease of use to allow for in-house expansion of the LUDA base file to include:

- . point data, e.g. well records, environmental sensor stations;
- . line data, e.g. roads, rivers, utility rights-of-way;
- . incremented details and revisions, as they become available.

Sample acquisition costs:

. digitizing table - "Bendix" (42" x 60", incl interfaces and x-y coordinate display)	$16,500#
. keyboard and storage tube CRT (for file editing) - "Tectronix"	9.500
. software - "Geodats"	40,000##
COST:	$66,000

\# Costs given to nearest $500. Trademarks (in quotes) are given for items tentatively selected as "optimum" among those researched. Where market range exceeds 20% of item cost, representative prices are given.

\## Estimated partial of multi-function package.

II. Interactive Geographic Analysis (IGA)

A. In this context, "interactive" does not simply mean some specified response time, but refers to a particular frame of mind relative to the task at hand. The goal is a fundamentally heuristic mode of analysis, with the computer used as the sounding board for a gradually refined understanding. In some situations, response time will be the critical factor; in others, it may be more important to preserve a record of each step in the analytic process itself for further study. Accordingly, the IGA subsystem must provide both 'hard' and 'soft' media of interaction - both cathode ray tube images and markable maps -- which in turn implies moth batch and real-time modes of computation.

B. The term "geographic analysis" assumes the integration of a conventional data base management system (DBMS) with a geographic information system (GIS). The DBMS handles the logical or structural relationships among data items, while the GIS deals with their spatial distribution. Both forms of inquiry must be available in the twin analytic modes distinguished above.

C. These two paris of requirements yield four IGA "application situations" for which hardware and software implications can be specified:

- "soft" (real-time) DBMS inquiry
  Hw: alphanumeric CRT terminal, on-line storage
  Sw: data base creation/maintenance language

- "soft" (real-time) GIS inquiry
  Hw: vectoring CRT terminal w/lightpen or cursor, display processor with DMA channel, more storage
  Sw: spatial search, thematic mapping, logical operators for file combination, legend supplies

- "hard" (batch-mode) DBMS inquiry
  Hw: teletype terminal, line printer (COM?)
  Sw: report generation language

- "hard" (batch-mode) GIS inquiry
  Hw: data entry terminal, plotter (COM?)
  Sw: automatic thematic mapping packages

D. Assuming flat storage and a teletype will be provided by the Dedicated Support Processor (DSP) as described below, the following cost estimates can be made for the implementation of the IGA subsystem:

HARDWARE

- Terminal + interface - "TTY ASR33" (proven durability chosen here over newer, less expensive -- but untried -- models) — $2,500

- 17" vectoring CRT (refreshed), w/light pen, display processor (DMA) and interfaces — 10,000 (x)#

  Note: refreshed CRTs demand more of c.p.u. but are the only way to distort an image in real time. Also, they are potentially less expensive than a storage-tube CRT of equal size--one of which is included with digitizer

- 20" electrostatic printer/plotter, inc. device software and interface - "Versatec" — 15,500

  Note: prints at 300lpm, plots at .75ips regardless of plot complexity; can produce true halftone shading (e.g. for SYMAP)

- COM (Computer Output Microfilm); to be purchased as a service when needed. Vendor picks up 800 bpi magnetic tape, delivers film or prints. — ---

SOFTWARE

- DBMS - as described above — $500 - $10,000

  Note: DOT has neither the core space nor the available storage capacity to support this function, which must therefore be handled by the DSP in-house. The price range reflects different "bundling" policies among vendors.

- GIS - as described above — 45,000

  Note: this appears to be a unique product. It includes the digitizing/editing software broken out in section I above.

#Areas where industry search is incomplete are marked (x). If a price is given, it represents the lowest price yet found for an adequate unit.

. Batch mapping packages (SYMAP, CALFORM, GRIDS and POLYVRT have already been purchased; the Census Bureau will provide a package for producing full color analytic maps on a COM).	---
. Report Generation language and statistical packages (already supported at DOT data center)	---
TOTAL IGA COSTS	$73,500

III. Dedicated Support Processor (DSP)

A. Five major functions must be supported concurrently by the DSP:

- Digitizing/editing of the GBF
- Real-time graphic analysis using the refreshed CRT
- DBMS inquiry and maintenace
- Communications interface with DOT batch-processing (CMS/RJE)
- Development of new applications programs

In addition, primarily through the medium of industry compatible magnetic tape, the DSP will provide a link to various outside service organizations, such as digitizing and COM contractors.

B. Preliminary estimates, based on the experiences of USGS, the Census and DCA/DSRP, project a GBF of some 10 million characters--depending on the chosen scope and resolution of the file. The DBMS must, in turn, assume a similar size. In an environment of heuristic problem analysis, each new task can be expected to generate many new files--data overlays, policy models with alternate projections, etc. Even with the highest hopes for good housekeeping among the files, it seems only prudent to configure the system with considerable elbow room in its mass storage. And in a geographic analysis sytem, that must mean on-line storage.

C. Considering the processor functions and storage requirements, the following estimates can be made about the costs of an adequate DSP:

- CPU (central processing unit) with 48k of main memory and a real time executive, including console, cabinet, and necessary options $43,000(x)

  Note: the minicomputer industry is extremely competitive. Marketing techniques make it hard to configure identical systems for price comparison. To date, three manufacturers have indicated an ability to meet these performance requirements: Data General, Digital Equipment Corp., and Hewlett-Packard.

.	Magnetic Disk storage (80M bytes)	$32,000 (x)
	The same comments apply here.	
.	Magnetic Tape drive - 2400ft, 800bpi	10,500
.	RJE teleprocessing link - 7200 Baud, leased line	3,000
.	Binary input/output - for system diagnostics (either paper tape or magtape cassetts)	1.500 (x)
	TOTAL DSP COST	$90,000

Appendix F:

IV. Staff Requirements

A. The system here proposed is intended as a planning tool, an aid in the fundamental process of problem definition and analysis, the development of a strategy for solution, the execution of that strategy and its inevitable refinement as the problem is better understood, and finally the documentation of work done as it relates to past and future problems. But it will not solve problems by itself.

B. Whatever specific tasks are given highest priority in the course of the Coastal Zone Management Program, a certain structural similarity will persist in how they are approached. That is inherent in the system. In this very general context, it may be useful to anticipate the personnel requirements implicit in the commitment to develop such a system.

C. Tasks and Costs

- A team of 3-4 planner/analysts @ $14-18K ea.
  - specify problem definition
  - develop and test alternate analytic approaches
  - document project results
- 2-3 programmers @ $12-16K ea.
  - develop new applications modules
  - run problem sets
  - expand and maintain DBMS
  - system documentation
- 1-2 technicians @ $8-12K ea.
  - digitizing and GBF maintenance
  - system library maintenance
- Hardware maintenance contracts @ $10K

TOTAL STAFF COSTS: 100K/yr (approx).

TOTAL SYSTEM COSTS

I.	GBF	$66,000
II.	IGA	73,500
III.	DSP	87,000
	Purchase Price#:	$229,500

#Many details of financing are ignored here, e.g. government discount, lease-to-purchase arrangements, option for rental of some elements.